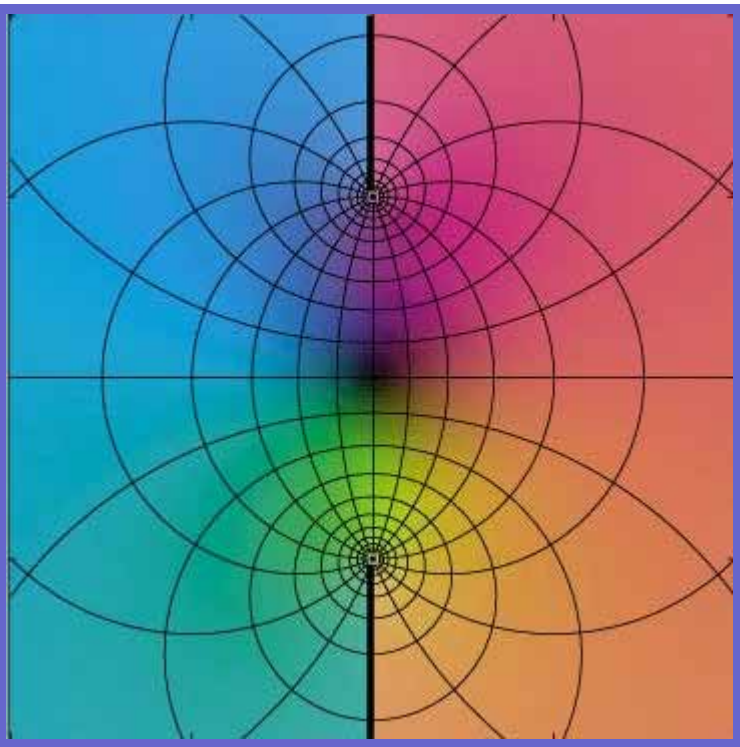
ISSN 0974 – 3340

SUTRA

International Journal of

Mathematical Science Education

VOLUME I, ISSUE I JULY-DECEMBER 2008



Published by

Technomathematics Research Foundation, India

Sutra: International Journal of Mathematical Science Education

EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India*.

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain*.
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK*.
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India*.
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India*.
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India*.
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India*.
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India*.
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia*.
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

Copyright ©2008 ***Technomthematics Research Foundation.***

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.
Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.
Sutra: International Journal of Mathematical Science Education (ISSN 0974 – 3340) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education

CONTENTS (Vol. 1, No. 1, 2008)

Editorial *Rupali Akerkar* iii

1. Mathematics Thrillerium *L. Radhakrishna.* 1 - 8
2. Accurate Solution Estimate and Asymptotic Behavior of Nonlinear Discrete System *Trupti P Shah and Raju K George.* 9 - 15
3. On Divisibility Tests And The Curriculum Dilemma *Yaakov Nahir.* 16 - 29
4. Class-room Notes *R. Akerkar.* 30 - 35
5. Mathematical Puzzle. 36

* Cover graphics: complex function atan(z) using Mathlab™

Sutra: International Journal of Mathematical Science Education

EDITORIAL (Vol. 1, No. 1, 2008)

It is with great pleasure we present this inaugural issue of *Sutra* to our readers. This journal is being published by the Technomathematics Research Foundation, India as a part of its initiatives to contribute constructively to mathematical science education in schools and universities. We hope through Sutra to reach researchers, students and teachers of mathematical sciences all over the world.

We would like to target audience from under-graduate classes and their teachers of mathematics. However, we do intend to include material of interest to graduate students and researchers in mathematical sciences. The Sutra issue will be published twice a year and we hope to include variety of contributions: research articles, general articles, series devoted to different aspects of a chosen topic in mathematical sciences, classroom notes for teachers and students, book reviews, to name a few.

We welcome contributions, comments ad suggestions from our readers. Our constant endeavour will be to enhance the quality and accessibility of material to our readers.

Our editorial board members, referees and colleagues have given us academic and moral support. We thank them all and express the hope that they and our contributors and readers will continually keep in touch with us and support this effort.

Rupali Akerkar

Editorial Coordinator

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 1, No. 1, 1 – 8, 2008

MATHEMATICS THRILLERIUM

L. Radhakrishna
UGC Centre for Advanced Studies in Fluid Mechanics
Dept of Mathematics, Bangalore University, Bangalore 560001, India
lankaradhakrishna@yahoo.co.in

The following nine thrills are discussed: The chaos in mathematics communication, World War III and mathematics, definition of science/art from the 21st century definition of mathematics, the paralysis in analysis, 'EinSTein' and self orthogonal vectors, the Celerity monoid, 'EINstein' as a correction to 'NEWton' and Glashow's snake, 'EINSTein' and the Taj Mahal of Science, the mathematician's mathematician and the Strange Indian Melody.

1 INTRODUCTION

From 00 to 97: The prestige of Mathematics Education has been raised to the prestige of Mathematics Research only in the 21st century— on the eve of World Mathematics Year 2000. To be specific, the esteemed journal *Mathematical Reviews* published by the American Mathematical Society has given the Primary Subject Classification Number 97 to the discipline Mathematical Education in their revised Subject Classification Scheme of 2000; prior to 2000, education had no independent status as a primary branch of enquiry and was classified under the Secondary Subject Classification number 00A35 (00 for General). From 00 to 97 is a thrilling jump in status for research in mathematics education. Recently a computer journal has put it thus: *Education is business. Customer (student) is the king*.

The *SUTRA* Initiative*:* A perfunctory professor— who does his teaching work as a means of earning his livelihood— confines himself to the prescribed syllabus. He has no aptitude for creating an excited feeling for the subject in his audience. It needs effort to pep up interest in the class room with

- History of the concepts
- Philosophy of the reasoning in proofs.

However a professional professor takes care of minute details of presentation besides history and philosophy of the subject, so as to ensure a wonder-joy session; that is, start the lecture with wonder—a captivating caption— and end the lecture with joy due to the thrill of understanding. This is desirable but difficult and so perfunctory professors outnumber the professional ones. It is a tribute to the sagacity of the founders of the new journal *SUTRA*, for focusing on the educational aspects. The aim of this article, is to help the perfunctory teacher to evolve into a professional one via thrillerium—a typical collection of philosophical/historical thrills about mathematics. For instance one can replace the matter-of-fact title "Galois Group of a Field" in a class room, by the arresting caption "The Climax in Elegance in Algebra". It raises the attention curve of the students. We present in this small article some topics to add spice to pedagogy.

2 THE FIRST THRILL: The Chaos in Mathematics Communication

The Birth and Growth of Mathematics: Who invented the word 'mathematics'? Since 'history of mathematics' is not included in the syllabus, it is not a popular question in educational institutions, the said question becomes a quiz. The answer will help in appreciating how ancient the subject is. In fact it is the oldest discipline, since the word 'mathematics'—that which can be learnt— was invented by Pythagoras (584-495 B.C.). He coined the word

'philosophy' as well. Mathematics is 25 centuries old, while physical sciences are about one hundred years old. Biology, Statistics and Psychology are hardly 70 years old. The following table gives an estimate of the growth of the discipline in question:

Year	Number of Primary Subjects	Number of Secondary Subjects
500B.C.(Pythagoras era)	2	4
1868 A.D.	12	38
1979 A.D.	60	3400
2000 A.D.	67	14000

In the year 1980 about 200,000 theorems were published in 1600 journals [Davis and Hersh 1981]. The doubling period for mathematics being 10 years, in 2008 we safely estimate 800,000 theorems per annum in 14000 branches of mathematics [see the table above]. To verify the veracity of one theorem, a professional can take one day or even one month. This is an exponential proliferation.

The Chaos of Communication: It is not possible for one person to keep track of all the research in the 14000 specializations. With increasing technical jargon, the chaos in communication even among mathematicians is sufficiently complex! Consequently, how one can meaningfully convince the Caesars (the practitioners of the royal science of politics, who fund mathematics) — with the relevance and thrills of the subject— is much more challenging. A journal of education like *SUTRA* can provide partial order in the total chaos. Senior professionals may have to consider it a professional obligation to enlighten the funding agencies on the effectiveness of mathematics.

World War III and Mathematics: Science and technology provide an inexhaustible demand for more mathematical models. Linguistics and cliometry from the arts accelerate the thrill in the mathematization of culture. *Mathematics by fiat* is the clarion call from social welfare. From January 2000 Mathematical Reviews has introduced at Serial No.92 the major subject of research "Biology and other Natural Sciences". The successful mapping of the Human Genome Project in 2000A.D. is a tribute to mathematical techniques and theoretical computer science. Michael Atiyah eulogizes this ubiquitous role of mathematics thus: *The 20th century has transformed mathematics from a cottage industry run by a few semi-amateurs into a world-wide industry run by an army of professionals*" [P. viii, Arnold 2000] This army of professional mathematicians will play a major role in World War III, by monitoring the speed and accuracy of the delivery of weapon systems through efficient algorithms. Mathematics is the essence of algorithms— the soul of computers. It may be noted that World War I is credited to chemistry through trinitrotoluene (the gun powder) and World War II is credited to physics via release of atomic energy. Mathematics is more basic and ancient than physics/ chemistry and no wonder if it takes over World War III [Davis and Hersh 1981]. The education of such a prime discipline must be nurtured with priority. This is the purpose of SUTRA.

3 THE SECOND THRILL: The 21st Century Definition of Mathematics/ Science/Arts

"I love this definition…" declared the Fields Medalist David Mumford, in his address *Mathematics towards the Third Millennium* [P.199, Arnold 2000]. Since the particular description is popularized in the year 2000, we christen it as the 21st century definition, originally due to Davis and Hersh [P. 399, 1981]. It states that:

The study of mental objects with reproducible properties is called mathematics.

The thrills that follow from this definition are (i) if we replace the word *mental* by the word *physical*, we get the description of science. (ii) Keeping *mental* unchanged, if we replace *reproducible* by the word *irreproducible*, then we get the definition of arts.

Illustration: The elements $x, y, z, x * y, y * z$ of a set S, with the binary operation $*: S \times S \to S$, are mental objects, since x, y, z are all meaningless symbols. In calculus the functions $f, f', f'', \ldots$ with dashes representing derivatives are devoid of any meaning and so the symbols are mental objects. Since paintings and music are peculiar to its creator, the arts are not reproducible and so are subjective. Thus mathematics shares objectivity with science and it shares the study of nonphysical objects with the arts. However it is a paradox that mathematics *per se* is excluded from the Nobel Prize award unlike the sciences and the arts.

Note: Current dictionaries give definitions that are outdated by 4 centuries! For instance the Oxford dictionary (2003) cites the 16th century definition of mathematics:

The science of number, quantity, and space is called mathematics.

4 THE THIRD THRILL: THE PARALYSIS IN ANALYSIS

There are two perspectives on mathematics:

- The main stream approach
- The foundational approach.

In Section 3, we have followed the first perspective resulting in the latest definition of mathematics. The corresponding philosophy is named by Hersh as *Mathematical Humanism* [Grinstein and Lipsey 2001]. In this section we offer an exposition of the second approach. Foundations of mathematics deal with the laws of reasoning:

[R1] The Law of Excluded Middle: Every statement is either true or false.

[R2] The Law of Trichotomy: Every real number is either positive or negative or zero.

[R3] The Modified Law of Excluded Middle: Every statement is either true or false or

neither true nor false.

For describing the philosophies, we need the thrilling result:

[GP] Gödel's Proposition: Mathematics is fallible.

Here we recall the famous Incompleteness theorem due to Gödel which states that 'there is no algorithm whose output contains all true statements of arithmetic and no false ones' [P.311 Hersh 1997]. There are three beliefs on the subject:

[M1] Mathematics is invented.

[M2} Mathematics is discovered.

[M3] Mathematics that can be constructed in a finite number of steps is genuine.

Now we can describe the three warring groups owing allegiance to certain standard philosophies:

[P1] *Formalism* characterized by [R1], [R2] and [M1] with Hilbert as the exponent.

[P2] *Mathematical Platonism* featured by [R1], [R2], and [M2] with Einstein in the front

[P3] *Constructivism or Intuitionism* satisfying [R3], [GP] and [M3] with Brouwer as leader.

They respectively characterize pure mathematics, applied mathematics and computer science.

While Russell has demonstrated the falsity of [R1] through the construction of R-set, Brouwer has demonstrated the falsity of [R2] by constructing 'pi hat' [Davis and Hersh 1981]. It is well known that Cantor's theorem on the uncountability of real numbers is established by *reductio ad absurdum*, which is an indirect proof. Such proofs are based on the two valued logic of [R1], which is false for infinite sets. Besides the construction of real number system is not possible in a finite number of steps. This violates [M3]. Hence real analysis is not genuine mathematics according to [P3]. It is unfair mathematics as it uses the reasoning [R1], which is not true. Till a direct proof of the uncountability of real numbers is discovered by some genius in future, real analysis remains paralyzed. It follows that complex analysis and functional analysis share the paralysis at present. This is responsible for the birth of constructive functional analysis, circumventing the use of real numbers.

5 THE FOURTH THRILL: ''EinSTein" and the Self Orthogonal Vectors

The icon of the millennium: Washington based weekly *TIME* conducted an 'on line' poll from January 1999 to December 1999, to tell the world the Person of the 20th century and announced that Albert Einstein lead all the rest since he 'revised the universe with an equation'. The magazine identified the persons of each of the centuries from 11th to 20th (For instance Queen Elizabeth for 16th century, Sir Isaac Newton for 17th century) and compared their contributions to art, science, society etc. Again Einstein's influence was the most. Thus Einstein (1879-1955) became *the icon of the second millennium* (1001-2000) [Issacson *2000*]. We attempt a scintillating portrait of the 'most creative intellect of the human race' Einstein, by splitting the word 'Einstein' in three different ways and associate an appropriate mathematical thrill with each split.

"Ein ST ein": Einstein was a German and in German language 'ein' means 'one'. As the word Einstein starts with 'ein' and ends with 'ein', we infer that the name 'Ein ST ein' suggests that he obtained the unification or one-oneness of S and T. What do S and T in the middle could stand for? From history of mechanics, the obvious inference is that S refers to 'space' and T refers to 'time'! This unification of space-time was achieved in 1905 by Einstein. This amalgamation of S and T into ST is enshrined in his name 'EinSTein' itself. What a divine dispensation! From 1905 onwards 'events' form the basic data of physics and not just 'points'. Einstein's 'Interval'

$$\tau = \sqrt{(-x^2 - y^2 - z^2 + c^2t^2)} \tag{5.1}$$

between the events (0, 0, 0, 0) and (x, y, z, ct) where c is the velocity of light, became the focus of investigation rather than Euclid's 'Distance'

$$s = \sqrt{(x^2 + y^2 + z^2)} \tag{5.2}$$

between the points (0, 0, 0) and (x, y, z).

Self-orthogonal 4-vectors: From (5.2), we have

$$s^2 = 0 \text{ if and only if } x = 0, y = 0, z = 0, \tag{5.3}$$

but from (5.1) we deduce

$$\tau^2 = 0 \text{ if and only if } x^2 + y^2 + z^2 = c^2t^2. \tag{5.4}$$

For example the 4-vector $v = (x, 0, 0, ict)$ with $i^2 = -1$, is orthogonal to itself, since

$$v.v = (x^2 + 0^2 + 0^2 + i^2c^2t^2)$$

$$= 0, \text{ by (5.4).}$$

The inner product of v with itself vanishes and so it is self orthogonal. The concept of self orthogonal vectors is against common sense and so it is thrilling. The tangent to the light ray is self orthogonal.

6 THE FIFTH THRILL: EINSTEIN'S CELERITY MONOID

Let $(R, +, \cdot)$ denote the field of real numbers with usual addition, multiplication and c be a fixed privileged real number. The letter c is the first letter of the word '*celerity*' (speed). We define Einstein's composition (of high speeds)

$\oplus: R \times R \to R$

by $$(x, y) \to x \oplus y = (x + y) + (1 + (x \cdot y)c^{-2}). \qquad (5.5)$$

The composition $\oplus$ is a binary operation, since $x \oplus y$ is in R. We observe that

$$(x \oplus y) \oplus z = x \oplus (y \oplus z) \quad \forall x, y, z \in R \quad \text{(Associativity property)}$$

and $$x \oplus 0 = x = 0 \oplus x \qquad \forall x \in \quad \text{(Existence of identity).}$$

Hence we infer that $(R, \oplus)$ is a monoid.

Note the thrilling properties involving *c*:

(i) $c \oplus (-c) = 0 + 0$, which is indeterminate. (5.6)

(ii) The equation $c \oplus x = 0$ has no solution in x since it implies the untenable result $c = 0$; that is, the inverse of the privileged element *c* does not exist. So the monoid does not qualify to be a group.

(iii) $c \oplus x = c \qquad \forall x \in R.$

(iv) $c \oplus c = c.$ (5.7)

Bisociation: We recall a relation in real analysis analogous to (5.7). If *c* is the cardinality of the continuum, then [Simmons 1963]

$$c \blacksquare c = c \qquad (5.8)$$

where $\blacksquare$ represents the cardinal addition. Eq. (5.8) is often expressed as $\aleph_1 + \aleph_1 = \aleph_1$, with $\aleph_1$ as the uncountable actual infinity of real numbers. Such unexpected formal similarity in Equations (5.7) and (5.8) in different disciplines — high speed mechanics and real analysis—are referred to as 'bisociation', which is expected to be the dominant feature of 21st century mathematical research [Browder 2002]. In the algebra of electrical circuits and in Boolean algebra we have a similar result

$$x + x = x, \quad \forall x.$$

Remark 1: The equation (5.5) is the familiar relation in Einstein's special relativity for the composition of collinear velocities comparable to the velocity of light.

Remark 2: In Einstein's result (5.7), *c* is finite but huge equal to 1, 86,000 miles per hour, while in Cantor's result (5.8), c is the uncountable infinity. What a contrast in the scope of *c*!

Remark 3: With the potential infinity ∞ and the real number x the properties (iii) and (iv) are analogous to

$$\infty + x = \infty \text{ and } \infty + \infty = \infty.$$

7 THE SIXTH THRILL: 'EINstein' IS A CORRECTION TO 'NEWton'

In German 'ein' means 'one' and 'stein' means a 'stone' and a stone as a measure stands for 14 pounds. In England the new ton by weight stands for 2240 pounds. Obviously 14 is a correction (zero error) to 2240. This is the literal picture when we say: 'EIN stein' is a correction to 'NEW ton'. This thrilling interpretation of the two legendary names can be justified mathematically in more than one way. [Metaphorically also the contribution of Newton is more fundamental than Einstein. Without Newton's differential calculus of the 17th century, Einstein's tensor calculus of the 20th century would not have been possible.] Specifically to enlighten the small correction to Newton's formulae, we give two illustrations.

[1] Newton's famous formula for the force $\boldsymbol{F}$

$$\boldsymbol{F} = m\boldsymbol{a},$$

gets Einstein's modification for high speeds

$$\boldsymbol{F} = m\boldsymbol{a} + (\boldsymbol{F} \cdot \boldsymbol{u})\boldsymbol{u}c^{-2}.$$

The second term on the right hand side of this equation is negligible when u is small compared to c.

[2] In Newtonian mechanics, the planet Mercury describes an elliptical orbit given by the differential equation

$$D^2 u + u = mh^{-2}$$

where u is the reciprocal of the radial distance of the planet from the sun and D is the time derivative, m is the mass and h is a certain constant. Careful experiments have shown that Mercury's orbit was not a perfect ellipse. Einstein gave the corrected equation:

$$D^2 \text{ u} + \text{u} = \text{m}h^{-2} + 3mu^2c^{-2}.$$

The minuteness of the correction can be gauged, when it (the perihelion motion of Mercury) is found to be ***17 seconds of arc per century.*** One has to wait for a hundred years to see a slight deviation from elliptical orbit for the fastest planet Mercury.

8 THRILL SEVEN: 'EINstein' and GLASHOW'S SNAKE

The normal weight of the head of a person is 14 pounds or one stone and so the word 'EINstein' for 'one stone' can mean 'the one head' or 'the unique brain'. The head of Einstein has become photographers' delight, with his tousled mane and furrowed forehead. Encyclopedia Britannica (1960 edition) described Einstein as 'the most creative intellect of the human race'. Who will be remembered in the year 3001 A.D.? Certainly Einstein, since atom bomb, the Big Bang picture of the universe, and electronics, television, space travel bear his stamp.

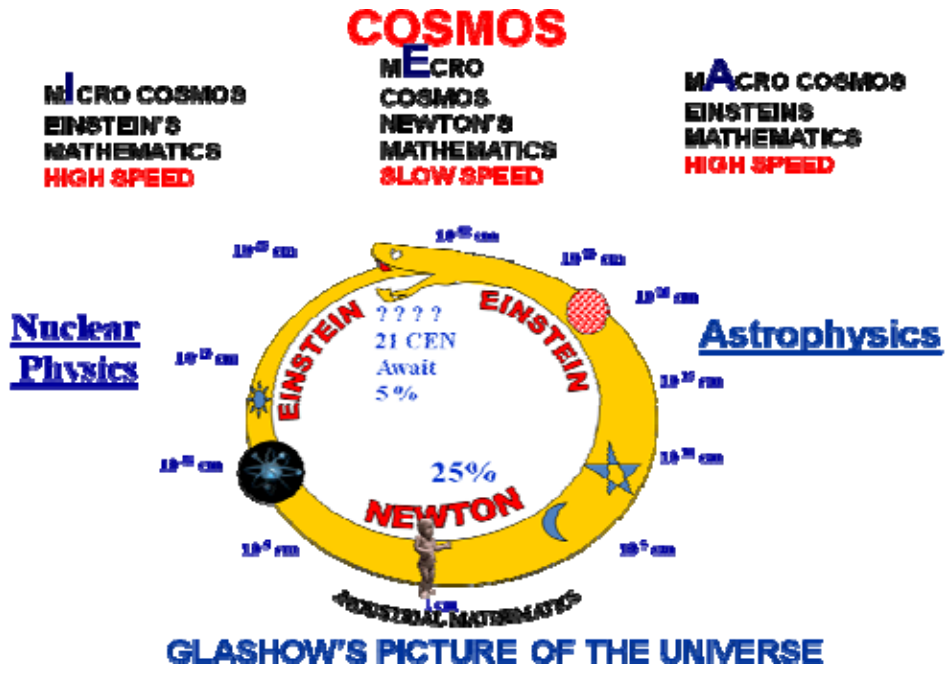
Einstein coined the words tensor, photon, and cosmos. Micro cosmos relevant to nuclear physics involves high speeds and so Einstein's mathematics of Special Relativity is relevant. Macro cosmos deals with super massive bodies at high speeds and at high pressures and the appropriate mathematical tool is Einstein's General Relativity. In the figure below, conceived by the Nobel laureate Glashow, micro cosmos occupies 35% of the universe and macro cosmos occupies 35%. It means 70% of the universe can be studied with Einstein's mathematics. For such a prolific individual, God must have designed the name in a special way!

In contrast, the slow speed mechanics of Newton is applicable to 25% of the cosmos, sometimes referred to as mecrocosmos— the universe accessible to the naked eye. This implies that mathematics of only 5% of the cosmos (the common portion between the hood and the tail of the snake *ouroborus*) remains to be discovered. In fact Einstein's theory fails at

Planck length 10^{-33}centimeters and at Planck time 10^{-44} seconds.

What happens below the Planck length and Planck time is a challenge for the 21st century scientists.

String theory is struggling to come to the rescue. We have to wait for an Einstein-like genius! They are so rare!



9 EIGHTH THRILL: 'EINSTein' AND TAJ MAHAL OF SCIENCE

The combination of the German words 'Einst' and 'ein' can be interpreted as 'the first one'. This refers to 'the first' discipline created entirely by just 'one man'. In other words the first person to create a subject without history in science is Einstein. The subject is General Relativity. In fact his happiest thought was [Narlikar 1978]

Gravitation ≡ Curvature of space-time.

He then announced the field equations of gravitating matter referred to as Taj Mahal of Science

(vide *Gravity* by George Gamow):

$$R_\beta^\alpha - \frac{1}{2} R \delta_\beta^\alpha = -8\pi G c^{-4} T_\beta^\alpha .$$

The left hand side is a divergence free tensor (known as the Einstein tensor) representing the geometry which is as beautiful and strong as the marble of the Taj Mahal. The right hand side is the stress-energy-momentum tensor of gravitating matter and π, G, c are usual constants. Consequently Einstein predicted in 1916 the then thrills:

(i) Light bends in the neighborhood of ponderable masses (verified by experiment in 1921)

(ii) Universe is expanding (verified by Hubble in 1919)

(iii) Planets have quirks in their elliptic orbits (not understood before 1916).

The experimental verification of Einstein's mathematical predictions on the existence of gravitational waves and Black holes is inconclusive to this day.

10 NINTH THRILL: THE MATHEMATICIAN'S MATHEMATICIAN RAMANUJAN

A scientist becomes immortal in the history of mathematics just by announcing one conjecture. Ramanujan (1887-1920) has written about 600 conjectures, some of which are still being checked for validity. This is a record and such natural mathematical talents are almost rare. Newman has called him 'The Mathematician's Mathematician' as distinguished from 'Poincare: The Physical Mathematician' [Newman 1972]. An exemplar of his genius eulogized as 'Strange Indian Melody' (an infinite continued fraction) by Hofstadter [1983] is reproduced below:

$$\frac{1}{1+}\,\frac{\exp(-2\pi\sqrt{5})}{1+}\,\frac{\exp(-4\pi\sqrt{5})}{1+}\cdots \text{to } \infty = \left(\sqrt{5}\,[1+A]^{-1} - B\right)\exp(2\pi/\sqrt{5})$$

where

$$A = \sqrt[5]{5^{3/4}\,\Omega - 1}\,,\quad B = \frac{(\sqrt{5}+1)}{2}\,,\quad ? = \left(\sqrt{5-1}/2\right)^{5/2}.$$

"He was a mathematician so great that his name transcends jealousies, the one superlatively great mathematician whom India produced in the last thousand years" [Kanigel 2003].

CONCLUSION

In this paper a sample of nine thrills have been designed out of author's teaching and research experience. The historical anecdotes add value to mathematical pedagogy. The philosophical discourses provide a unifying command of exposition. Every teacher must have experienced thrills while reading or explaining to students. These exciting facts could be shared for the benefit of future practitioners through the columns of *SUTRA*.

REFERENCES

Arnold et al (Ed) *Mathematics: Frontiers and Prospects,* American Mathematical Society, 2000.

Browder, F., Reflections on the Future of Mathematics, *Notices of the American Mathematical Society*, **49,** 658-663, 2002.

Davis, P.J. and Hersh R., *The Mathematical Experience*, Cambridge Birkhauser 1981.

Hersh, R., *What is Mathematics Really?* Oxford University Press, 1997.

Hersh, R 'Philosophical Perspectives in Mathematics Education' *in* Grinstein, L.S., and LipseyS.I. (Eds.) *Encyclopedia of Mathematics Education,* Routledge Falmer 2001.

Hofstadter, D.R., *Gödel, Escher, Bach: An Eternal Golden Braid,* Penguin Books, P.563, 1983.

Issacson, W. 'Person of the Century' in Knaeur (Ed) *Time 100,* Time Books, Volume III, 82-89, 2000.

Kanigel R., *The Man Who Knew Infinity: A Life of the Genius Ramanujan,* Abacus 2003.

Narlikar J. V., *Lectures on General Relativity and Cosmology*, Mac Millan (India), 1978.

Newman, E. 'Srinivasa Ramanujan' in Klein M. (Ed) *Mathematics in the Modern World: Readings from Scientific American,* Freeman, 1972.

Simmons G.F. Introduction to Topology and Modern Analysis, Mc Graw Hill, 1963.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 1, No. 1,9-15, 2008

ACCURATE SOLUTION ESTIMATE AND ASYMPTOTIC BEHAVIOR OF NONLINEAR DISCRETE SYSTEM

TRUPTI P SHAH

Department of Applied Mathematics,
Faculty of Technology & Engineering,
M.S.University of Baroda,
Vadodara 390001, India.
tpshah-appmath@msubaroda.ac.in

RAJU K GEORGE

Department of Mathematics,
Indian Institute of Space Science and Technology,
Trivandrum - 695 022. India.
rkg.iist@gmail.com

This article deals with the nonlinear discrete dynamical system

$$x(t+1) = A(t)x(t) + f(t, x(t)), \quad t \in N_0 = \{0, 1, 2, ...\}$$

We first derive accurate estimate for the norm of solution of this system. This gives us stability condition and bound for the region of attraction of the stationary solution. We also give sufficient conditions for the asymptotic stability of the null solution of the above system. Our approach is based on the concept of generalized subradius for the coefficient matrices. Numerical example showing asymptotic behavior of the null solution is also given to support our result.

Keywords: generalized subradius ; asymptotic stability; discrete Gronwall's inequality.

1. Introduction

Let R^n be the set of n-real vectors endowed with the Euclidean norm $\| \cdot \|$. Consider the n dimensional discrete system

$$x(t+1) = A(t)x(t) + f(t, x(t)), \quad t \in N_0 = \{0, 1, 2, ...\} \tag{1}$$

where $(A(t))_{t \in N_0}$ is a sequence of real $n \times n$ matrices and $\{f(t, x(t)\}_{t \in N_0}$ is a sequence of nonlinear continuous vector functions.

Medina and Gil derived accurate estimates for the norms of solutions using the approach based on "freezing" method for difference equations and on recent estimates for the powers of a constant matrix (see [Medina and Gill (2004)]). Also many authors have studied the asymptotic stability of the null solution of such systems.

A well-known result of Perron which dates back to 1929 (see [Gordon (1972)], [Ortega (1973)] , page 270 and [LaSalle (1976)] , Theorem 9.14) states that system (1) is asymptotically stable with $A(t) \equiv A$ (constant matrix) provided that spectral radius $\rho(A)$ of A is less than 1 and $f(t,x) = o(\| x \|)$. We established asymptotic stability of the null solution using a concept of (sp) matrix and taking some growth condition on f (see [George, Shah (to appear)])

To obtain the estimates for the norms of solution and asymptotic stability of the null solution of system (1), we use the recent concept of generalized subradius.

Czornik introduced the the ideas of generalized spectral subradius and the joint spectral subradius and shown the relationship between generalized spectral radii and the stability of discrete time-varying linear system (refer [Czornik (2005)])

Let Σ denote a nonempty set of real $n \times n$ matrices. For $m \geq 1$, Σ^m is the set of all products of matrices in Σ of length m,

$$\Sigma^m = \{A_1 A_2 ... A_m : A_i \in \Sigma, \;\; i = 1, 2, ..., m\}$$

Denote by $\rho(A)$ the spectral radius and by $\| A \|$ a matrix norm of the matrix A. Let $A \in \Sigma^m$.

Definition 1.1. *The generalized spectral subradius of* Σ *is defined as*

$$\tilde{\rho_*}(\Sigma) = \inf_{m \geq 1} (\inf_{A \in \Sigma^m} \rho(A))^{\frac{1}{m}}.$$

Definition 1.2. *The joint spectral subradius of* Σ *is defined as*

$$\hat{\rho}_*(\Sigma) = \inf_{m \geq 1} (\inf_{A \in \Sigma^m} \| A \|)^{\frac{1}{m}}.$$

Czornik proved that the generalized spectral subradius and joint spectral subradius are equal for any nonempty set Σ and the common value of $\tilde{\rho_*}(\Sigma)$ and $\hat{\rho}_*(\Sigma)$ is called the generalized subradius of Σ and will be denoted by $\rho_*(\Sigma)$ (see [Czornik (2005)]).

Let

$$\Phi(t,0) = A(t-1)A(t-2)...A(0), \;\; \Phi(t,t) = I$$

denotes the fundamental matrix of the system

$$x(t+1) = A(t)x(t) \quad , t \in N_0 \tag{2}$$

and

$$x(t, 0, x_0) = \Phi(t,0)x_0$$

is the unique solution of equation (2) with initial condition $x(0, 0, x_0) = x_0$.
In [Czornik (2005)], the following result is proved.

Proposition 1.3. *([Czornik (2005)]) Consider the discrete time-varying linear system (2). where $(A(t))$ is a sequence of matrices taken from Σ. Then there exists a sequence $(A(t))$ such that for any initial state $x_0 \in R^n$, we have $\lim_{t\to\infty} x(t) = 0$ if and only if $\rho_*(\Sigma) < 1$.*

Note that if $\rho_*(\Sigma) < 1$, then for fixed $d \in (\rho_*(\Sigma), 1)$ there exist matrices $A_{i_1}, A_{i_2}, ..., A_{i_{n_0}} \in \Sigma$ such that $\| A_{i_1} A_{i_2} ... A_{i_{n_0}} \| < d$.

We recall that the zero solution of (1) is stable if for any $\epsilon > 0$ there exist $\delta > 0$ and $t_0 \in N_0$ such that

$$\| x(t) \| < \epsilon$$

whenever $t \geq t_0$ and $\| x_0 \| < \delta$. We say that the zero solution of (1) is absorbing if for any $x_0 \in R^n$,

$$\lim_{t\to\infty} \| x(t) \| = 0.$$

We say that the zero solution of (1) is asymptotically stable if it is both stable and absorbing.
We extend the above result for the nonlinear system (1) under the following assumptions.

Assumptions :

(a) Let $\rho_*(\Sigma) < 1$ and

$$\| \Phi(t, j) \| \leq \beta_j, \quad for \ \ all \ \ 0 \leq j \leq t < \infty.$$

where $\sum_{j=0}^{\infty} \beta_j \leq \eta$ for some constant $\eta > 0$.

(b) There exist constants $q, \mu \geq 0$ such that

$$\| f(t, x) \| \leq q \| x \| + \mu, \ \ \forall x \in B_r = \{x \in R^n : \| x \| < r\} \ \ for \ \ some \ \ r > 0.$$

(c) The constants q and η defined above satisfy $q\eta < 1$.

2. Main Results

Now we derive accurate estimate for the norm of solution of system (1).

Theorem 2.1. *Under Assumptions (a), (b) and (c) any solution $\{x(t)\}_{t=0}^{\infty}$ of (1) satisfies the inequality*

$$\sup_{t\geq 0} \| x(t) \| \leq \frac{\beta_0 \| x_0 \| + \mu\eta}{1 - q\eta}$$

provided that

$$\| x_0 \| \leq \frac{r(1 - q\eta) - \mu\eta}{\beta_0} \ \ for \ \ some \ \ r > 0$$

Proof. By inductive arguments, it is easy to see that the unique solution $\{x(t)\}_{t=0}^{\infty}$ of (1) under initial condition $x(0) = x_0$ is given by

$$x(t) = \Phi(t,0)x_0 + \sum_{j=0}^{t-1} \Phi(t,j+1)f(j,x(j))$$

i.e.

$$\| x(t) \| \leq \| \Phi(t,0) \| \| x_0 \| + \sum_{j=0}^{t-1} \| \Phi(t,j+1) \| (q \| x(j) \| + \mu) \quad if \quad \| x \| < r$$

Hence,

$$\| x(t) \| \leq \beta_0 \| x_0 \| + q \max_{i=0,1,\ldots t-1} \| x(i) \| \sum_{j=0}^{t-1} \| \Phi(t,j+1) \| + \mu \sum_{j=0}^{t-1} \| \Phi(t,j+1) \|$$

i.e.

$$\| x(t) \| \leq \beta_0 \| x_0 \| + q \max_{i=0,1,\ldots t-1} \| x(i) \| \sum_{j=0}^{\infty} \| \Phi(t,j+1) \| + \mu \sum_{j=0}^{\infty} \| \Phi(t,j+1) \|$$

i.e.

$$\max_{i=0,1,\ldots t} \| x(i) \| \leq \beta_0 \| x_0 \| + q\eta \max_{i=0,1,\ldots t} \| x(i) \| + \mu\eta$$

$$\sup_{i \geq 0} \| x(i) \| (1 - q\eta) \leq \beta_0 \| x_0 \| + \mu\eta$$

i.e.

$$\sup_{i \geq 0} \| x(i) \| \leq \frac{\beta_0 \| x_0 \| + \mu\eta}{(1 - q\eta)}$$

provided that

$$\| x_0 \| \leq \frac{r(1 - q\eta) - \mu\eta}{\beta_0}$$

Hence the theorem. □

We now prove the asymptotic stability of the null solution of system (1).

Theorem 2.2. *Suppose that the system (1)satisfies*

(i) $lim_{t \to \infty} \sum_{j=0}^{t-1} \| \Phi(t,j+1) \| = 0$
(ii) $\rho_*(\Sigma) < 1$ *and*
(iii) Assumption (b).

Then the zero solution of (1) is asymptotically stable.

Proof. The unique solution $\{x(t)\}_{t=0}^{\infty}$ of (1) under initial condition $x(0) = x_0$ is given by

$$x(t) = \Phi(t,0)x_0 + \sum_{j=0}^{t-1} \Phi(t,j+1)f(j,x(j))$$

Hence

$$\begin{aligned} \| x(t) \| &\leq \| \Phi(t,0) \| \| x_0 \| + \sum_{j=0}^{t-1} \| \Phi(t,j+1) \| \| f(j,x(j)) \| \\ &\leq \| \Phi(t,0) \| \| x_0 \| + \sum_{j=0}^{t-1} \| \Phi(t,j+1) \| (q \| x(j) \| + \mu) \\ \textit{i.e.} \| x(t) \| &\leq \{\| \Phi(t,0) \| \| x_0 \| + \mu \sum_{j=0}^{t-1} \| \Phi(t,j+1) \|\} + \sum_{j=0}^{t-1} q \| \Phi(t,j+1) \| \| x(j) \| \end{aligned}$$

Now applying the discrete Gronwall's inequality,

$$\| x(t) \| \leq \{\| \Phi(t,0) \| \| x_0 \| + \sum_{j=0}^{t-1} \mu \| \Phi(t,j+1) \|\} \prod_{j=0}^{t-1} [1 + q \| \Phi(t,j+1) \|]$$

Since condition (ii) and Proposition 1.3 implies that

$$\| \Phi(t,0) \| \| x_0 \| \to 0, \quad as \quad t \to \infty$$

Hence by using condition (i), we prove

$$\| x(t) \| \to 0 \quad as \quad t \to \infty$$

Hence the proof. □

3. Numerical Example

Example 3.1. *Consider the non-autonomous system given by the following equation.*

$$x(t+1) = A(t)x(t) + f(t,x(t)) \tag{3}$$

where the sequence $A(t) = \begin{pmatrix} \frac{1}{3}t & 0 \\ 1 & .04t \end{pmatrix}$, *Let* $\Sigma = \{A(h) : h = 0.002, 0.004, .006, ...\}$ *and*
$f = \frac{1}{6}\begin{pmatrix} \frac{1}{2}x_2(t) + \epsilon x_2^2(t) \\ \frac{1}{2}x_1(t) \end{pmatrix}$. *Let* $\epsilon \in (0, \frac{1}{2})$. *For* $n = 10$, *we can verify that the*

generalized spectral radius $\rho_*(\Sigma) = 0.008 < 1$ *and* $\eta = 1$. *Also*

$$\begin{aligned} \| f(t,x) \|^2 &= \frac{1}{36}\{\frac{1}{4}x_2^2 + \epsilon x_2^2 + \epsilon^2 x_2^4 + \frac{1}{4}x_1^2\} \\ &\leq \frac{1}{36}\{\frac{1}{4}x_2^2 + \epsilon x_2^2 + \epsilon^2 x_2^2 + \frac{1}{4}x_1^2\} \\ &\leq \frac{1}{36}\{(\frac{1}{2}+\epsilon)^2 x_2^2 + \frac{1}{4}x_1^2\} \\ &\leq \frac{1}{36}(\frac{1}{2}+\epsilon)^2 \| x \|^2 \\ &\leq \frac{1}{36}d^2 \| x \|^2 \ taking \ \ d = (\frac{1}{2}+\epsilon) < 1 \\ i.e., \| f(t,x) \| &\leq \frac{d}{6} \| x \| \end{aligned}$$

$$\| f(t,x) \| \leq q \| x \| \quad taking \ \ q = \frac{d}{6}$$

i.e. f *satisfies the Assumption* (b) *with* $q = \frac{d}{6} = \frac{0.7}{6} = 0.1167 < 1$ *taking* $\epsilon = 0.2$ *without loss of generality and* $\mu = 0$. *The function* f *is continuous and it satisfies* $f(0) = 0$. *Note that* $q\eta = 0.1167 < 1$. *Hence all the Assumptions of Theorem (2.2) are satisfied. We can see in Figure* 1 *that initial state* $x_0 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$ *reaches to null state in* 10 *iterations. Hence the null solution of equation (3) is asymptotically stable.*

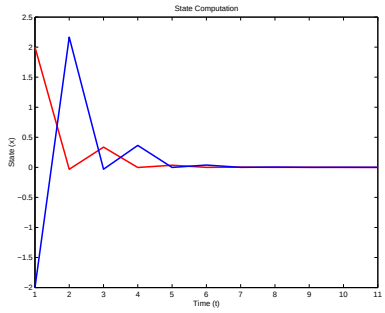


Fig. 1.

References

Czornik A. (2005). *On the generalized spectral subradius*, Linear Algebra and its Applications, **5**, 407 : 242-248.

Elaydi, S. N. (1996). *An Introduction to Difference equations*, Springer-Verlag New York, Inc.

George,R.K., Shah, T.P. *Asymptotic Stability of Nonlinear Discrete Dynamical Systems Involving (sp) Matrix*, (to appear in Nonlinear Studies).

Gordon, S.P. (1972). *On converses to the stability theorems for difference equations*, SIAM J. Control **10**: 76-81.

Kelley, W.G.,Peterson, A.C. (2001). *Difference equations*, Academic Press.

LaSalle, J.P. (1976). *The Stability of Dynamical Systems, Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathemtics, Pennsylvania.*

Medina, R., Gil, M. (2004). *Accurate solution estimates for nonlinear nonautonomous vector difference equations*, Abstract and Applied Analysis **7**: 603-611.

Ogata, K. (1995). *Discrete-Time Control Systems*, Second Edition, Pearson Education, Inc.

Ortega, J.M. (1973). *Stability of difference equations and convergence of iterative processes* SIAM J. Numer. Anal. **10**: no. 2, 268-282.

Xue, X., Guo, L. (2007). *A kind of nonnegative matrices and its application on the stability of discrete dynamical systems* , J. Math. Anal. Appl., Vol. **331**: Issue 2, 1113-1121.

Sutra:International Journal of Mathematical Science Education
©Technomathematics Research Foundation
Vol. 1, No. 1, 16-29, 2008

ON DIVISIBILITY TESTS AND THE CURRICULUM DILEMMA

Yaakov Nahir
Department of Mathematics, Ben-Gurion University
Beer-Sheva 84105, Israel
eisen@013.net

Abstract

Constructing rules for determining the divisibility of whole numbers is a subject that is being dropped from the curriculum. Few teachers today know if a simple division of whole numbers will give an integer quotient without doing the actual division, nor do they have the skills to construct such a test. This paper takes the stance that this is a shame, and it tries to rectify the situation by presenting several different methods for constructing rules of divisibility. Some of the methods presented are known but not well-known, while others are completely new; yet all are within the grasp of elementary school teachers.

Keywords: Divisibility rules, divisibility criteria, whole number division, proof.

1. Background: two sides to the argument. Calculators and computers have changed the school mathematics curriculum to the point where even experienced teachers are pondering the merits and demerits of drill and practice ($D\&P$), a notion that used to stand at the heart of the curriculum for "cementing in" ideas. Mathematics educators have been debating the role of $D\&P$ for years, with each side giving passionate arguments as to why their thinking on the subject should be adopted by curriculum decision makers, and by those who are still uncertain as to which side of the polemic to endorse. Do $D\&P$ really help students develop a deeper understanding and appreciation for the notions under study, or is it as the adversaries claim, "boring to the students," and a major contributing factor as to why students hate mathematics? The answer to this will never be definitive within our profession as a whole, but each teacher must answer this question for themself. Where does one draw the line these days with respect to $D\&P$, taking into account the existence of sophisticated computer algebra systems where most of the problems encountered in school mathematics can be solved in nanoseconds, if one knows how to set up the computer to solve them?

Every teacher has lists of essential topics and skills they believe all children should know, but unfortunately, the topics on these lists and the depth of the knowledge we wish to impart to the children concerning these topics are not standard, even with respect to the simplest of notions. E.g., should children be expected to know how to multiply a three digit whole number by a two digit whole number? I, and thousands of other teachers, say "yes"; but I am certain that just as many teachers

can be found supporting a negative answer on this topic–and they hold to this negative stance even when the issue is phrased in a personal way: do you want your children (or grandchildren) to be able to correctly carry out the work long-hand to compute (538)(79)? To my dismay, I have colleagues who claim that they do not care whether or not their own children can carry out such a multiplication. They validate this stance by saying that we are living in the 21^{st} century, where calculators and computers are everywhere; they can be found even on one's wristwatch. And what about $\frac{1}{2} + \frac{1}{3}$? (They seem not to care about this either!)

2. A rationale for studying rules of divisibility. Probably the main reasons for studying rules of divisibility are to help students experience the thrill of creating mathematics, and to give one a sense of intellectual accomplishment. I doubt if anyone these days is ever going to calculate whether or not 7 divides 7568940320156 by hand to see if one obtains a whole number quotient. But developing the machinery to answer this question is another matter.

Listed below are a few different ways to know if $\frac{a}{b} = c$, where a, b, and c are whole numbers ($b \neq 0$). The methods presented in sections 3 and 4 are known but not well known; however I believe the rules presented in section 5 are completely new.

3. Rules of divisibility. Most students can tell if a natural number is divisible by $2, 3, 4, 5, 6, 8, 9$, and 10. For completeness, we list the rules below and make brief comments on some of them.

- **For 2:** A natural number $N = abcde$ is divisible by 2 if its last digit is even.
- **For 3:** A natural number $N = abcde$ is divisible by 3 if the sum of its digits is divisible by 3. E.g., $3 \Big| (a+b+c+c+d+e)$. This is easy to see by writing the number N in expanded form:

$$\begin{aligned} N &= abcde \\ &= a(10^4) + b(10^3) + c(10^2) + d(10) + e \\ N &= a(9999+1) + b(999+1) + c(99+1) + d(9+1) + e \\ &= \underbrace{a \cdot 9999 + b \cdot 999 + c \cdot 99 + d \cdot 9} + \underbrace{a + b + c + d + e} \end{aligned}$$

It is well known that for whole numbers N, m and n where $N = m + n$ and for prime number p, if p divides N, and p divides m, then p also divides n. I.e.,

$$\begin{aligned} N &= m + n \\ p \Big| N \text{ and } p \Big| m &\longrightarrow p \Big| n. \end{aligned}$$

There are methods to rigorously prove this statement, but students readily

accept it in an intuitive manner without proof. So in summary,

$$If\ 3\Big|N \text{ and } 3\Big|\underbrace{a\cdot 9999+b\cdot 999+c\cdot 99+d\cdot 9}, \text{ then } 3\Big|\underbrace{a+b+c+d+e}.$$

I.e., if 3 divides N, then 3 divides the sum of the digits in N, and vice versa.

- **For 4:** A natural number $N = abcde$ is divisible by 4 if the number formed by the last two digits of N is divisible by 4. E.g., if $4\Big|de$, then $4\Big|N$.
  This is also easy to see by writing the number N in a slightly different form.

$$\begin{aligned} N &= abcde \\ &= a(10^4)+b(10^3)+c(10^2)+d(10)+e \\ &= \underbrace{a(10^4)+b(10^3)+c(10^2)}+\underbrace{d(10)+e} \end{aligned}$$

  The number 4 divides all terms on the right hand side of the above equality except the two digit number de. So if 4 divides the two digit number formed by de, then the number N is divisible by 4.
- **For 5:** A natural number $N = abcde$ is divisible by 5 if the last digit of N is either a 5 or a zero. This rule is immediately obvious to students and it is generally accepted without elaboration.
- **For 6:** Here $N = abcde$ is divisible by 6 if N is divisible by both 2 and 3. I.e.

$$\text{If } 2\Big|N \text{and if } 3\Big|N, \text{then } 6\Big|N.$$

  Again, there are rigorous ways to prove this statement, but students immediately accept the rule as being obvious.
- **For 7:** The test for 7 is a bit tricky, so we'll look at shortly.
- **For 8:** Here $N = abcde$ is divisible by 8 if the number formed by the last three digits of N is divisible by 8. That is, if the number cde is divisible by 8. The proof is very similar to the one for divisible by 4.

$$\begin{aligned} N &= abcde \\ &= a(10^4)+b(10^3)+c(10^2)+d(10)+e \\ &= \underbrace{a(10^4)+b(10^3)}+\underbrace{c(10^2)+d(10)+e} \end{aligned}$$

  The number $10^4 = 10000$ is divisible by 8 and so then is the number $a(10^4)$. Similarly for the number $b(10^3)$. But the number 10^2 is not divisible by 8. So writing N as

$$N = abcde$$

$$= a(10^4) + b(10^3) + c(10^2) + d(10) + e$$

$$= \underbrace{a(10^4) + b(10^3)} + \underbrace{c(10^2) + d(10) + e},$$

we see that the number 8 divides all terms on the right hand side of the equality above except the three digit number formed by *cde*. So if 8 divides the three digit number formed by *cde*, then the number N is divisible by 8. It should also be noted that since $8 \Big| a(10^4) + b(10^3)$, the number 8 will also divide the larger numbers

$$x(10^5) + a(10^4) + b(10^3),$$

$$y(10^6) + x(10^5) + a(10^4) + b(10^3), \text{ etc.}$$

So no matter how large the whole number is (when written in base 10), it will be divisible by 8 if the number formed by the last three digits of it is divisible by 8.

- **For 9:** The rule for divisibility by 9 is the same as it is for divisibility by 3, and it's proof follows immediately from it. A number N is divisible by 9 if the number formed by the sum of its digits is divisible by 9.

Back to divisibility by 7.

Method 1: The procedure developed here follows from the above by writing the whole number N in expanded form. Let

$$\begin{aligned} N &= abcdefgh \\ &= a10^7 + b10^6 + c10^5 + d10^4 + e10^3 + f10^2 + g10 + h \\ &= a\Big(1428571(7) + 3\Big) + b\Big(142857(7) + 1\Big) + c\Big(14285(7) + 5\Big) + d\Big(1428(7) + 4\Big) \\ &\quad + e\Big(142(7) + 6\Big) + f\Big(14(7) + 2\Big) + g\Big(1(7) + 3\Big) + h(1) \end{aligned}$$

Rearranging gives

$$\begin{aligned} N &= abcdefgh \\ &= a10^7 + b10^6 + c10^5 + d10^4 + e10^3 + f10^2 + g10 + h \\ &= \underbrace{a\Big(1428571(7)\Big) + b\Big(142857(7)\Big) + c\Big(14285(7)\Big) + d\Big(1428(7)\Big) + e\Big(142(7)\Big) + f\Big(14(7)\Big) + g\Big(1(7)\Big)} \\ &\quad \underbrace{+ a\Big(3\Big) + b\Big(1\Big) + c\Big(5\Big) + d\Big(4\Big) + e\Big(6\Big) + f\Big(2\Big) + g\Big(3\Big) + h\Big(1\Big)} \end{aligned}$$

So, if $7 \Big| N$, then 7 must divide

$$\underbrace{a\Big(3\Big) + b\Big(1\Big) + c\Big(5\Big) + d\Big(4\Big) + e\Big(6\Big) + f\Big(2\Big) + g\Big(3\Big) + h\Big(1\Big)}.$$

We now do a trick: If $7\Big|4c$, then $7\Big|(-3)(4c)$. And if $7\Big|(-3)(4c)$ then $7\Big|(4)(-3c)$. But since 7 does not divide 4, then we know that 7 **must divide** $(-3c)$. Ie.,

$$7\Big|4c \text{ implies that } 7\Big|(-3\text{c}).$$

This "trick" goes under the name of "modular" arithmetic. We say $a \equiv b(\text{mod p})$ implies "a" when divided by "p" leaves the same remainder as when "b" is divided by "p". This is equivalent to saying that "p" divides "(a-b)". So with respect to the above: $4 \equiv -3(\text{mod } 7)$.

Returning to the above, 7 must divide $\underbrace{a\Big(3\Big)+b\Big(1\Big)+c\Big(5\Big)+d\Big(4\Big)+e\Big(6\Big)+f\Big(2\Big)+g\Big(3\Big)+h\Big(1\Big)}$ is equivalent to saying that 7 must divide

$$\underbrace{a\Big(3\Big)+b\Big(1\Big)+c\Big(-2\Big)+d\Big(-3\Big)+e\Big(-1\Big)+f\Big(2\Big)+g\Big(3\Big)+h\Big(1\Big)}.$$

$$\underbrace{\Big\{a(3)+b(1)\Big\}-\Big\{\text{c}(2)+\text{d}(3)+\text{e}(1)\Big\}+\Big\{\text{f}(2)+\text{g}(3)+\text{h}(1)\Big\}}.$$

We have just developed a test for divisibility by 7. We take a number $N = abcdefgh$ and starting on the right hand side, we block off the digits into groups of three.

$$\begin{aligned} N &= abcdefgh \\ &= ab\{cde\}\{fgh\} \end{aligned}$$

Starting with the units digit and moving from right to left, we multiply the units digit by 1, the digit in the tens place by 3, and the digit in the hundreds place by 2. We continue on doing this from right to left; we multiply the digit in thousands place by 1, the digit in the ten-thousands place by 3 and in the one-hundred thousands place by 2 and we continue on moving from right to left. We then take the algebraic sum as indicated by the formula we have developed. If this algebraic sum is divisible by 7, then the number N is divisible by 7. And if this algebraic sum is not divisible by 7, then N is not divisible by 7.

E.g., does $7\Big|56798435492$? Does 7 divide

$$\{3(5)+1(6)\}+\{2(7)+3(9)+1(8)\}-\{2(4)+3(3)+1((5)\}+\{2(4)+3(9)+1(2)\}?$$

If so, then $7\Big|56798435492$.

This method of finding "a set of magic multipliers" for the digits in the blocked off number will always work. But although it works beautifully for developing a test

for divisibility by 11, in general, it isn't very efficient, particularly for large prime divisors. Luckily for us, there is a more efficient method for deciding divisibility by 7 and other low-valued primes.

4. An efficient procedure. Let's again look at divisibility by 7. Suppose we want to know if $7 \Big| N = abcdef$. Chopping off the units digit of N gives us the number $M = abcde$. So multiplying M by 10 and the adding on the units digit returns us to N. Ie.,

$$\begin{aligned} N &= abcdef \\ M &= abcde \\ 10M &= abcde0 \\ 10M + f &= abcde0 + f = N. \end{aligned}$$

Now we do a trick:

$$\begin{aligned} N &= abcdef \\ N &= abcde0 + f \\ N &= abcde0 + (21f - 20f) \\ N &= 10\Big(abcde - 2f\Big) + 21f. \end{aligned}$$

Now if 7 divides N then 7 divides the right hand side of the last equation above. The number 7 most certainly divides $21f$, and so 7 must also divide $10\Big(abcde - 2f\Big)$. But 7 does not divide 10, so this means that 7 must divide $\Big(abcde - 2f\Big)$.

We have just developed a procedure for determining if $7 \Big|$ N. We chop off the units digit of N, multiply it by 2 and form the new number $abcde - 2f$. If 7 divides the new number, then 7 divides N; and if 7 does not divide $abcde - 2f$, then 7 does not divide N.

If N is large, then repeat the procedure. E.g. Does $7 \Big|$ 56789341?

$$7 \Big| 56789341 \ ?$$

$$7 \Big| 5678932 \ ?$$

$$7 \Big| 567889 \ ?$$

$$7 \Big| 56770 \ ?$$

$$7 \Big| 5677 \ ?$$

$$7 \Big| 553 \ ?$$

$$7 \Big| 49 \text{ ? yes!}$$

Therefore, $7 \Big| 56789341$.

This method of chopping off the units digit and then "playing" with the number as above always works, particularly with low valued primes, and students seldom have trouble with discovering divisibility rules by themselves. Here is another example: Find a procedure to determine if N is divisibly by 23.

Assume $N = abcdef$ then

$$\begin{aligned} N &= abcdef \\ M &= abcde \\ 10M &= abcde0 \\ 10M + f &= abcde0 + f \\ N &= 10M + f = abcde0 + f \\ N &= abcde0 + \big(70f - 3(23)f\big) \\ N &= 10(abcde + 7f) - 3(23)f. \end{aligned}$$

So if $23 \Big| N$ then 23 must divide the right hand side of the last equation above. The number 23 most certainly divides $-3(23)f$ so 23 must divide $10(abcde+7f)$. But 23 does not divide 10, so then 23 must divide $abcde + 7f$, and we have just developed a procedure for testing if a number N is divisible by 23.

Similar tests can be developed for low valued primes with unit digits of $\{1, 3, 7, 9\}$. Odd as it might seem, tests of divisibility for 7 and other low-valued primes are not well-known and few mathematics teachers seem to be able to construct them, even though Eisenberg (2000), Hoch (2001) and I, Nahir (2003), have tried to convince teachers to explicitly include them in their lessons–not for the rule itself, but for the thrill of discovering it. In this spirit, I now present an even more efficient way than what is presented in this section. I discovered these rules by musing on the above, and although they can be presented in an abstract manner, I present them here in a simplified way.

5. An even more efficient procedure for two-digit divisors.

5.1 For divisors ending in 1. Suppose that the units digit of a two digit divisor is 1 and that the units digit for N is u. We write the divisor d as $(10a + 1)$ and number N as $(10[N/(10)] + u)$, where $[XX]$ is the greatest integer function. Then $(10a + 1) \Big| N$ or

$$(10a + 1) \Big| (10[N/10] + u)$$

if, and only if, there is a number k such that

$$[N/10] - ua = k(10\text{a} + 1).$$

Proof: Suppose

$$\begin{aligned}
[N/10] - ua &= k(10a+1) \\
[N/10] &= ua + k(10a+1) \\
10[N/10] &= 10ua + 10k(10a+1) \\
10[N/10] + u &= 10ua + u + 10k(10a+1) \\
10[N/10] + u &= u(10a+1) + 10k(10a+1), \\
10[N/10] + u &= (10a+1)(u+10k) \text{ but} \\
10[N/10] + u &= N, \text{and so} \\
N &= (u+10k)(10a+1)
\end{aligned}$$

That is, N is divisible by $(10a+1)$ and the quotient is $q = (u+10k)$. Moreover, since each step in the above proof is reversible, the statement is indeed an if, and only if, theorem. And additionally, we obtain the quotient without doing the actual division!

Example 1: Does $61 \Big| 2257$? Here

$$\begin{aligned}
N &= 2257,\ u = 7,\ a = 6 \\
2257 &\rightarrow 225 - (7)(6) = 183 = k \cdot 61 \\
k &= 3 \text{ and thus, } 61 \Big| 2257.
\end{aligned}$$

To obtain the actual quotient q we work backwards:

$$q = (u+10k) = 7 + 10 \cdot 3 = 37, \text{ and indeed } 2257 = 37 \cdot 61.$$

Example 2: Does $71 \Big| 34293$? Since the expression $N - ua$ is not easily recognized as a multiple of the divisor $(10a+1)$, we use the procedure recursively as we did in Section 4. The idea is to a build a chain until we find an appropriate integer k that is a multiple of the divisor, and then work backwards to obtain the actual quotient.

$$\begin{aligned}
N &= 34293,\ u_0 = 3,\ a = 7 \\
34293 &\rightarrow 3429 - (3)(7) = 3408 = k_0 \cdot 71 \\
3408 &\rightarrow 340 - 8(7) = 284 = k_1 \cdot 71.
\end{aligned}$$

Here it is easy to see that $k_1 = 4$. So we know that $71 \Big| 34293$. To find the actual quotient q we simply calculate backwards.

$$\begin{aligned} k_1 &= 4, u_1 = 8, u_0 = 3 \\ k_0 &= u_1 + 10(k_1) = 8 + 40 = 48 \\ q &= u_0 + 10k_0 = 3 + 10(48) = 483. \text{ And indeed} \\ 34293 &= \Big(3 + 10(48)\Big)(71) = (483)(71). \end{aligned}$$

5.2 For divisors ending in 3. Suppose the two digit divisor is of the form $(10a+3)$. Then

$$(10a + 3) \Big| (10[N/10] + u)$$

if, and only if, there is a number k such that

$$[N/10] - u(7a + 2) = k(10\text{a} + 3).$$

Proof: Suppose

$$\begin{aligned} [N/10] - u(7a + 2) &= k(10a + 3) \\ [N/10] &= u(7a + 2) + k(10a + 3) \\ 10[N/10] &= 70ua + 20u + 10k(10a + 3) \\ 10[N/10] + u &= 70ua + 21u + 10k(10a + 3) \\ 10[N/10] + u &= 7u(10a + 3) + 10k(10a + 3), \\ 10[N/10] + u &= (10a + 3)(7u + 10k) \text{ but} \\ 10[N/10] + u &= N, \text{and so} \\ N &= (7u + 10k)(10a + 3) \end{aligned}$$

That is, N is divisible by $(10a + 3)$ and the quotient is $q = (7u + 10k)$.

Example 1: Does $43 \Big| 2881$? Here $u_0 = 1, a = 4$.

$$2881 \to 288 - (1)(30) = 258 = k_0 \cdot 43.$$

It is easy to see that $k_0 = 6$, and so $q = 7u_0 + 10k_0 = 67$, and $2881 = (67)(43)$.

Example 2. Does $53\Big|2575641$? Here $u_0 = 1, a = 5$. Using the greatest integer notation $[n]$ we see that

$$\begin{aligned}
N &\rightarrow [N/10] - u(7a+2) = k(10a+3) \\
2575641 &\rightarrow 257564 - (1)(37) = 25764 - 37 = 257527 = k_0(53) \\
257527 &\rightarrow 25752 - (7)(37) = 25752 - 259 = 25493 = k_1(53) \\
25493 &\rightarrow 2549 - (3)(37) = 2549 - 111 = 2438 = k_2(53) \\
2438 &\rightarrow 243 - 8(37) = 243 - 296 = -53 = k_3(53).
\end{aligned}$$

So $k_3 = -1$ and we know that $53\Big|2575641$. We obtain the actual quotient by computing the values of k_3, k_2, k_1, k_0 with resect to the various unit digits $u_0 = 1, u_1 = 7, u_2 = 3$ and $u_3 = 8$. Specifically,

$$\begin{aligned}
k_3 &= -1 \\
k_2 &= 7u_3 + 10k_3 = 7(8) + 10(-1) = 46 \\
k_1 &= 7u_2 + 10k_2 = 7(3) + 10(46) = 481 \\
k_0 &= 7u_1 + 10k_1 = 7(7) + 10(481) = 4859 \\
q &= 7u_0 + 10k_0 = 7(1) + 10(4859) = 48597
\end{aligned}$$

And indeed, $2575641 = 48597(53)$.

5.3 For divisors ending in 7. Suppose the two digit divisor is of the form $(10a+7)$. Then

$$(10a+3)\Big|(10[N/10] + u)$$

if, and only if, there is a number k such that

$$[N/10] - u(3a+2) = k(10a+7).$$

Proof: Suppose

$$\begin{aligned}
[N/10] &= u(3a+2) + k(10a+7) \\
[N/10] &= 30ua + 20u + 10k(10a+7) \\
10[N/10] + u &= 30ua + 21u + 10k(10a+7) \\
10[N/10] + u &= 3u(10a+7) + 10k(10a+7)
\end{aligned}$$

$$
\begin{aligned}
10[N/10] + u &= 3u(10a+7) + 10k(10a+7), \\
10[N/10] + u &= (10a+7)(3u+10k) \text{ but} \\
10[N/10] + u &= N, \text{and so} \\
N &= (3u+10k)(10a+7)
\end{aligned}
$$

That is, N is divisible by $(10a+7)$ and the quotient is $q = (3u+10k)$.

Example 1: Does $67 \Big| 871$? Here the units digit $u_0 = 1$. So,

$$871 \to 87 - (1)(20) \quad = \quad 67 = k_0(67)$$

Thus k=1 and the quotient $q = (3u_0 + 10k_0) = 3(1) + 10(1) = 13$. And indeed $871 = (13)(67)$.

Example 2. Does $87 \Big| 2980098$?

$$
\begin{aligned}
N &\to [N/10] - u(3a+2) = k(10a+7) \\
2980098 &\to 298009 - (8)(26) = 297801 = k_0(87)(26); u_0 = 8, k_0 = ? \\
297801 &\to 29780 - (1)(26) = 29754 = k_1(87); u_1 = 1, k_1 = ? \\
29754 &\to 2975 - (4)(26) = 2871 = k_2(87); u_2 = 4, k_2 = ? \\
2871 &\to 287 - 1(26) = 261 = k_3(87); u_3 = 1, k_3 = ? \\
261 &\to 26 - 1(26) = 0(87) = k_4 = 0; u_4 = 1, k_4 = 0.
\end{aligned}
$$

And to get the actual quotient we reverse the steps with the values of k_4, k_3, k_2, k_1, k_0 and with unit digits $u_0 = 8, u_1 = 1, u_2 = 4, u_3 = 1$, and $u_4 = 1$. Specifically,

$$
\begin{aligned}
N &\to [N/10] - u(3a+2) = k \\
k_3 &= 3u_4 + 0k_4 = 3(1) + 0 = 3 \\
k_2 &= 3u_3 + 10k_3 = 3(1) + 10(3) = 33 \\
k_1 &= 3u_2 + 10k_2 = 3(4) + 10(33) = 342 \\
k_0 &= 3u_1 + 10k_1 = 3(1) + 10(343) = 3423 \\
q &= 3u_0 + 10k_0 = 3(8) + 10(3423) = 34254
\end{aligned}
$$

And indeed, $2980098 = 34254(87)$.

5.4 For divisors ending in 9. Suppose the two digit divisor is of the form $(10a+9)$. Then

$$(10a+9)\Big|(10[N/10]+u)$$

if, and only if, there is a number k such that

$$[N/10]-u(9a+8)=k(10a+9).$$

Proof: Suppose

$$\begin{aligned}
[N/10] &= u(9a+8)+k(10a+9)\\
[N/10] &= 90ua+80u+10k(10a+9)\\
10[N/10]+u &= 90ua+81u+10k(10a+9)\\
10[N/10]+u &= 9u(10a+9)+10k(10a+9)\\
10[N/10]+u &= 9u(10a+9)+10k(10a+9),\\
10[N/10]+u &= (10a+9)(9u+10k) \text{ but}\\
10[N/10]+u &= N, \text{and so}\\
N &= (9u+10k)(10a+9)
\end{aligned}$$

That is, N is divisible by $(10a+9)$ and the quotient is $q=(9u+10k)$.

Example 1: Does $29\Big|1073$? Here the units digit $u_0=3$.

$$1073 \quad\rightarrow\quad 107-(3)(26)=k\cdot 29;$$

So, $u_0=3$ and $k_0=1$ which forces the quotient to $(9u+10k)=9(3)+10(1)=37$; and indeed $1073=37(29)$.

Example 2: Does $59\Big|2863329$? Here the units digit $u_0=9$.

$$\begin{aligned}
N &\rightarrow [N/10]-u(9a+8)=k(10a+9)\\
2863329 &\rightarrow 286332-(9)(53)=2863325-467=285855=k_0(59)\\
285855 &\rightarrow 25585-(5)(53)=28320=k_1(59)\\
28320 &\rightarrow 2832-(0)(53)=2832=k_2(59)\\
28328 &\rightarrow 283-2(53)=177=k_3(59).
\end{aligned}$$

and we see that $k_3 = 3$. Therefore, $59 \Big| 2863329$.

In getting the quotient we note that the unit digits are $u_3 = 2, u_2 = 0, u_1 = 5, u_0 = 9$. And now we proceed as before:

$$\begin{aligned} N &\rightarrow [N/10] - u(9a+8) = k(10a+9) \\ k_3 &= 3 \\ k_2 &= 9u_3 + 10k_3 = 9(2) + 10(3) = 48 \\ k_1 &= 9u_2 + 10k_2 = 9(0) + 10(48) = 480 \\ k_0 &= 9u_1 + 10k_1 = 9(5) + 10(480) = 4845 \\ q &= 9u_0 + 10k_0 = 9(9) + 10(48450) = 48531 \end{aligned}$$

And indeed, $2863329 = 48531(59)$.

6. In summary. In this article we have presented several different ways to test whether or not for whole numbers d and N, $d \neq 0$, if $d \Big| N$ gives a whole number quotient. Such a unit of study used to be part of the school curriculum, but it has all but disappeared from it, and this paper takes the stance that it shouldn't. Students can gain much from studying such a unit, because hidden in it are doors of opportunity for having students experience the thrill of having created mathematical procedures by themselves, for asking why one method is preferred over another, and how the methods they have developed can be generalized. Granted, there are many such exercises that can be included in the curriculum to accomplish these lofty goals, but this particular topic seems to be one of the nicest, and one which is readily available to most students. My goal for writing this paper will be realized if a few of the readers will try teaching this topic to their students. I think that the types of thinking involved with this unit are particularly nice–and that the benefits from studying such a unit will spill over into other parts of the curriculum.

Acknowledgments

I would like to thank Professors T. Eisenberg and D. Berend of the Department of Mathematics at Ben-Gurion University, and Dr. Tzila Shamir of the Ariel University Center of Smaria, for their comments and suggestions concerning this paper.

References

Eisenberg, T. (2000). *On divisibility by 7 and other low-valued primes*, International Journal for Mathematics Education in Science and Technology, **31**, 622-626.

Hatch, G. (2001). *Divisibility tests for low-valued primes and their use as generators*

of simple proofs, International Journal for Mathematics Education in Science and Technology, **32**, 721-726.

Nahir, Y. (2003). *Tests of divisibility*, International Journal for Mathematics Education in Science and Technology, **34**, 581-591.

Sutra: International Journal of Mathematical Science Education
©Technomathematics Research Foundation
Vol. 1, No. 1, 30 - 35, 2008

CLASS-ROOM NOTES:

OPTIMIZATION PROBLEM SOLVING - I

R. Akerkar

Technomathematics Research Foundation
Kolhapur, India
kpr_tmrf@bsnl.in

Purpose of these notes is to discuss the optimization problem. An optimization problem consists of finding the best (cheapest, heaviest, etc.) element in a large set P, called the feasible region and usually specified implicitly, where the quality of elements of the set are evaluated using a function *f(x)*, the objective function, usually something fairly simple. The element that minimizes (or maximizes) this function is said to be an optimal solution of the objective function at this element is the optimal value. The greedy approach is a technique in which an optimal solution is obtained by means of the set of choices. A greedy algorithm always makes the choice that looks best at that moment. In this lecture, we will discuss greedy approach. To determine whether the greedy approach will yield the optimal solution the basics of matroid theory are presented.

Keywords: optimization, greedy algorithm, matroid theory.

1. Introduction

Greedy approach is a technique in which an optimal solution is obtained by the means of the set of choices. The idea of greedy approach is to make a locally optimal choice, hoping that this choice will lead to a globally optimal solution. There are many problems for which greedy approach provides an optimal solution much more quickly than a dynamic programming approach would have done.

Greedy algorithms are algorithms, which follow the problem solving meta-heuristic of making the locally optimum choice at each stage with the hope of finding the global optimum.

While most solution algorithms make choices based on a global overview of all current and future possibilities aiming at reaching the single global optimum solution, greedy algorithms make choices that look best at that very moment. In other words, they make locally shortsighted choices believing that the choice will eventually lead to the global optimum solution.

In general, greedy algorithms have five pillars (*Wikipedia*):

- A candidate set, from which a solution is created,

- A selection function, which chooses the best candidate to be added to the solution,
- A feasibility function, that is used to determine if a candidate can be used to contribute to a solution,
- An objective function, which assigns a value to a solution, or a partial solution, and
- A solution function, which will indicate when we have discovered a complete solution.

In order to develop a greedy algorithm the following steps can be used.

- Determine the optimal substructure of the problem.
- Develop a recursive solution.
- Prove that at any stage of the recursion, one of the optimal choices is the greedy choice. Thus it is safe to make the greedy choice.
- Show that all but one of the subproblems induced by having made the greedy choice are empty.
- Develop a recursive algorithm that implements the greedy strategy
- Convert the recursive algorithm to an iterative algorithm.

However in practice we usually streamline the above steps, when designing a greedy algorithm.

These steps can be reduced to more general sequence of steps.

1. Cast the optimization problem as one in which a choice is made and we are left with one subproblem to solve.
2. Prove that there is always an optimal solution to the original problem that makes the greedy choice.
3. Show that, having made the greedy choice, the part that remains is a subproblem with the property that if we combine an optimal solution to the subproblem with the greedy choice made, an optimal solution to the problem will be reached.

There is no way in general that one can specify if a greedy algorithm will solve a particular optimization problem. However if the following properties can be demonstrated, then it is probable to use greedy algorithm.

- *Greedy-choice property*: a globally optimal solution can be arrived at by making a locally optimal greedy choice. That is, when we are considering which choice to make, we make the choice that looks best in the current problem, without considering results from subproblems.
- *Optimal substructure*: A problem exhibits optimal substructure if an optimal solution to the problem contains within it optimal solutions to subproblems.

When compared to algorithms that guarantee to yield a global optimum solution: greedy-choice property, greedy algorithms have several advantages. They are easier to implement, they require much less computing resources, they are much faster to execute. Their only disadvantage is that they don't reach the globally optimal solution consistently, since they usually don't operate exhaustively on all the data. Nevertheless they are useful because they are quick to think up and often give good approximations to the optimum. If a greedy algorithm can be proven to yield the global optimum for a given

problem class, it naturally becomes the method of choice. To determine whether the greedy approach will yield the optimal solution the *matroid theory* can be used. This theory covers many cases of particular interest and is being rapidly developed and extended to cover more applications.

Matroids are abstract structures and they support greedy algorithms. It turns out that the structures, for which the greedy algorithm does lead to an optimal solution, are the matroids. It is worth studying them, not only because it enables us to recognize when the greedy algorithm applies, but also because there exist fast algorithms for `intersections' of two different matroids.

2. Essentials of Matroid Theory

The name "matroid" suggests a structure related to a matrix. Hassler Whitney introduced matroids theory in 1935 to capture abstractly the essence of dependence in linear algebra and graph theory. Since then, it has been recognized that matroids arise naturally in combinational optimization and can be used as a framework for approaching a diverse variety of combinational problems.

Definition 1: Let S be a finite set and let I be nonempty family of subsets of S. Then the pair $M=(S; I)$ is called a matroid if it satisfies the following conditions:

(i) $Æ \in I$,

(ii) If $B \in I$ and $A \subseteq B$ then $A \in I$,

(iii) If $A \in I$ and $B \in I$ and $|A| < |B|$ then $A \cup \{x\} \in I$ for some $x \in B\text{-}A$. (Exchange property)

Further, we will see some more notions.

Given $M=(S; I)$ matroid, element $x \notin A$ is called an *extension of* $A \in I$ id x can be added to A while preserving independence.

If A is an independent subset in a matroid M, A is *maximal* if it is not contained in any larger independent subset of M.

Let $Y \subseteq X$. A subset B of Y is called a *basis* of Y if B is an inclusion wise maximal independent subset of B. That is, for any set $Z \in I$ with $B \subseteq Z \subseteq Y$ one has $Z = B$.

A base of M is a maximal independent subset of S.

The common cardinality of all bases is called the rank of the matroid. If I is the collection of forests in a connected graph $G = (V, E)$, then the bases of the matroid (E, I) are exactly the spanning trees in G.

Definition 2: $M=(S; I)$ matroid is *weighted* if there is an associate weight function ω that assigns a strictly positive weight $\omega(x)$ to each element $x \in S$. the weight function ω extends to subsets of S by summation. $\sum^{w(x)}$ for any $A \subseteq S$.

In given $M=(S; I)$ matroid, a subset A where $A \in I$ and A has maximum possible weight is called an *optimal subset of matroid*.

There are different examples of matroids. Most useful ones are matric matroids, graphic matroids, uniform matroids and vectorial matroids. We will now discuss graphic matroids with certain details.

Graphic matroid can be defined as $MG=(SG, IG)$ in terms of given undirected graph $G=(V,E)$ in the following manner.

- The set SG is defined to be E, the set of edges of G.
- If $A \subseteq E$, then $A \in IG$ if and only if A is acyclic. That means, a set of edges A is independent if and only if the sub graph $GA=(V, A)$ forms a forest.

Theorem 1: *If G=(V, E) is an undirected g raph, then MG=(SG, IG) is a matroid.*
Proof: Conditions (i) and (ii) are trivial. To see that condition (iii) holds, let's assume that $GA=(V, A)$ and $GB=(V, B)$ are forests of G and that $|B|>|A|$. From graph theory it follows that a forest having k edges contains exactly $|V|-k$ trees. Thus, forest GA contains $|V|-|A|$ trees and forest GB contains $|V|-|B|$ trees.

Since forest GB has fewer trees than forest GA, forest GB must contain some tree T whose vertices are in two different trees in forest GA. Since T is connected, it must contain an edge (u, v) such that vertices u and v are in two different trees in forest GA $\Rightarrow$ the edge (u, v) can be added to forest GA without creating cycle. So MG satisfies the exchange property and is matroid.

Theorem 2: *All maximal independent subsets in a matroid have the same size.*
Proof: Proof follows from exchange property of matroid definition. (*Self-study*!)

Theorem 3*: (Closure axioms) A function* $s : 2^S \to 2^S$ *is the closure operator of a matroid on S if and only if for any* $A,B \subseteq S$ *and* $x, y \in S$ *the following conditions hold.*
(S1) $A \subseteq sA$.
(S2) If $A \subseteq B$ *then* $sA \subseteq sB$.
(S3) $sA = ssA$.
(S4) If $y \notin A, y \in s(A \cup \{x\})$, *then* $x \in s(A \cup \{y\})$.

Theorem 4*: (Independency axioms) a collection I of subsets S is the set of independent sets of a matroid on S if and only if the following conditions hold.*
(I1) $\varnothing \in I$
(I2) If $A \subseteq B \in I$, *then* $A \in I$.
(I3) If $A, B \in I$ *and* $|A| < |B|$, *then* $A \cup \{y\} \in I$ *for some* $y \in B - A$.

Theorem 5 *(Base axioms) a non-empty collection B of subsets S is the set of bases of a matroid on S if and only if the following conditions hold.*
(B1) If $B1, B2 \in B$ *and* $x \in B1 - B2$ *then* $B1 - \{x\} \cup \{y\} \in B$ *for some* $y \in B2 - B1$.

Theorem 6 *(Base axioms) a non-empty collection B of subsets of S is the set of bases of a matroid on S if and only if the following conditions hold.*

(B2) If B1, B2 Î B and x Î B1 – B2, then B2 È {x – y} Î B for some y Î B2 - B1.

3 Greedy Algorithm for Weighted Matroid

There are many problems that can be formulated as problems of finding a maximum-weight independent subset in a weighted matroid for which greedy approach can be used to provide optimal solutions. Greedy algorithm takes as input a weighted matroid $M=(S, I)$ with an associated positive weight function ω, and it returns an optimal subset A.

The greedy algorithm consists of selecting $y1,\ldots, yr$ successively as follows. If $y1,\ldots, yk$ have been selected, choose $y \in S$ so that:

(1) $y \notin \{y1,\ldots, yk\}$ and $\{y1,\ldots, yk, y\} \in I$,

(2) $\omega(y)$ is as large as possible among all y satisfying (1).

We stop if no y satisfying (1) exist, that is, if $\{y1,\ldots, yk\}$ is a basis.

In pseudocode Greedy algorithm can be presented like this.

```
Algorithm GREEDY (M, w)
        1. sort the elements according to weights: ω (y1) ≥ ω
           (y2) ≥ . . . ≥ ω (ym) ;
        2. let A= ∅ ;
        3. for i = 1 to m do
        4. if A ∪ {yi}  ∈ I then A = A È {yi}  .
```

The time cost is $O(n \log n + nf(n))$.

This is greedy algorithm is exactly Kruskal's algorithm (only for Kruskal algorithm the elements should be sorted in increasing order according to weights.).

Theorem 7: *If M = (S, I) is a weighted matroid with weight function w, then GREEDY(M,w) returns an optimal subset.*

The proof of the theorem is based on three lemmas discussed below.

Lemma 1: Let x be the first element of S such that $\{x\}$ is independent, if any such x exists. If x exists, then there exists an optimal subset A of S that contains x.

Proof: If no such x exists, then the only independent subset is ∅ and we are done. Otherwise, let B be any nonempty optimal subset.

Assume that $x \ddot{I} B$, otherwise $A = B$ is well done.

1. $\forall y \in B$, we have $w(y) \leq w(x)$ because of the choice of x.

2. Construct A as follows:

(a) Let $A = \{x\}$, then A is independent.

(b) Repeatedly find new element of B than can be added to A until $|A| = |B|$, preserving the independence of A.

(c) Then, $A = B - \{y\} È \{x\}$ for some $y \in B$. And $w(A) = w(B) - w(y)+w(x) \geq w(B)$

So, A must be also optimal and $x \in A$.

***Lemma 2*:** For any matroid $M = (S, I)$, if $x \in S$ is an extension of some independent subset A of S, then x is also an extension of ∅.

Proof: Since x is an extension of A, we have that $A \cup \{x\}$ is independent. Since I is hereditary $\{x\}$ must be independent. Thus, x is an extension of $\emptyset$.

Corollary 1: Let $M = (S, I)$ be any matroid. If $x \in S$ is not an extension of $\emptyset$, then x is not an extension of any independent subset A.

Remark: This corollary says that any element that cannot be used immediately can never be used. So, GREEDY never makes error when it skips over any initial element in S that is not an extension of $\emptyset$

Lemma 3 *:* Let $x \in S$ be the first element chosen by GREEDY for the weighted matroid $M = (S, I)$. The remaining problem of finding a maximum-weight independent subset containing x reduces to finding a maximum-weight independent subset of the weight matroid $M' = (S', I')$, where

1. $S' = \{y \in S \mid \{x, y\} \in I\}$
2. $I' = \{B \subseteq S - \{x\} \mid B \cup \{x\} \in I\}$, and the weight function for M' is the weight function for M, restricted to S'.

Proof:

1. If A is any maximum-weight independent subset of M containing x, then $A' = A - \{x\}$ is an independent subset of M'.
2. Conversely, any independent subset A' of M' yields an independent subset $A = A' \cup \{x\}$ of M.
3. Since we have in both cases that $w(A) = w(A') + w(x)$, a maximum-weight solution in M containing x yields a maximum weight solution in M', and vice versa.

Proof of Theorem 7: *using Lemmas proven above we can conclude.*

(1) Any elements that are passed over initially can be forgotten forever. Once an element x is selected, there must be an optimal containing it.

(2) $B \in I'$ is independent if $B \cup \{x\} \in I$ is independent. The subsequent operation of GREEDY will find a maximum-weight independent subset for M', and the overall operation of GREEDY will find a maximum-weight independent subset for M.

We will present various applications in the next issue of *Sutra*.

References

Akerkar, R. A. and Akerkar, R. R. *Discrete Mathematics*, Pearson Education (India), 2004.

Baase, S., Gelder, A. *Computer Algorithms.* Addison Wesley, 2000.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. *Introduction to algorithms, 2nd edition*. The MIT Press, 2001.

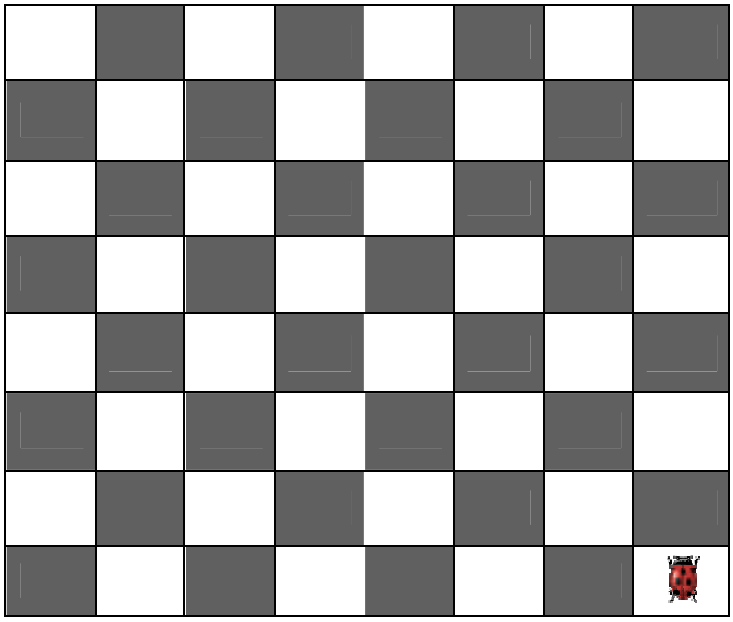
Oxley, J. G.: "Matroid Theory", Oxford University Press, New York, 1992.

Sutra: International Journal of Mathematical Science Education,
©Technomathematics Research Foundation
Vol. 1, No. 1, 36, 2008

Mathematical Puzzle

Readers are welcome to solve the following puzzles. But, do not send solutions to us. Answers to these puzzles will be provided in the next issue.

1. A bug begins from down right corner of a chess board and proceeds to visit every white square without ever entering a black square. If the insect never passes through the same intersection more than once, what should be his route?



2. Let us assume that we have given chess board with two of the diagonally opposite corners removed. Will it be viable to cover the chess board with pieces of tile whose size is exactly two board squares?

3. How many distinct 5-digit numerals can be created out of the digits 1; 1; 1; 6; 8?

4. Obtain the number of nonnegative integer solutions for the following equation

$$x_1 + x_2 + x_3 + x_4 < 19$$

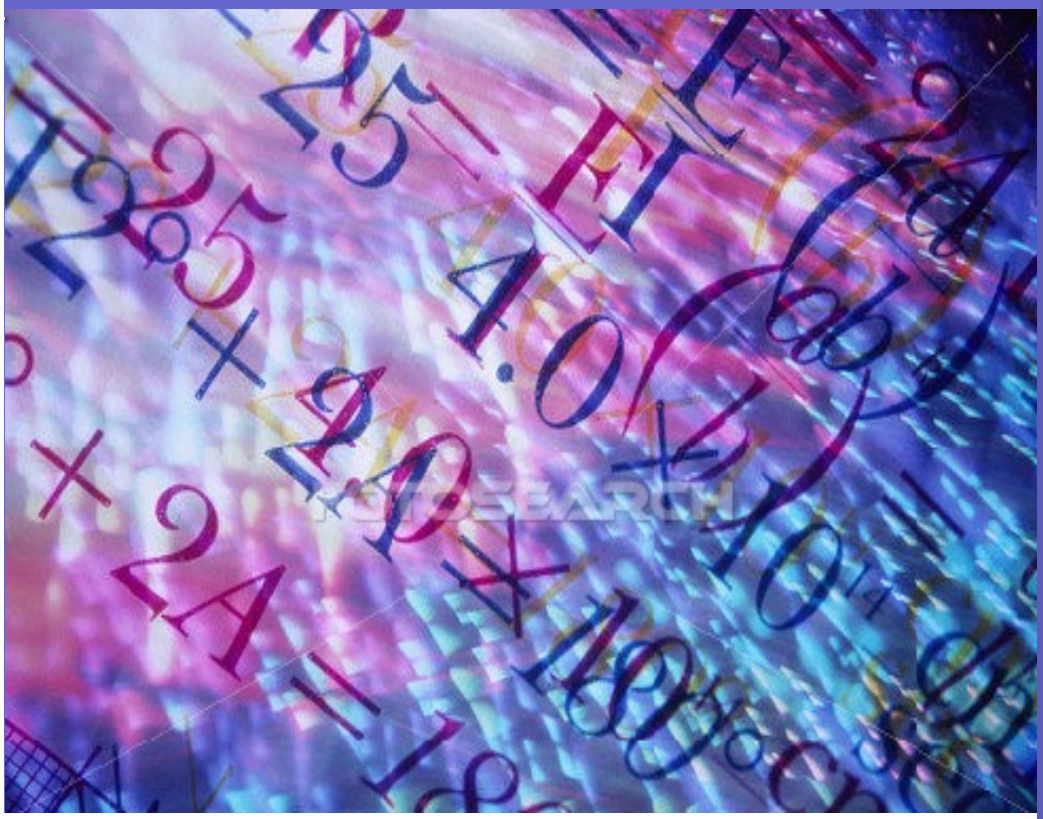
ISSN 0974 – 3340

S U T R A

International Journal of

Mathematical Science Education

VOLUME 2, ISSUE 1 JANUARY-JUNE 2009



Published by

Technomathematics Research Foundation, India

Sutra: International Journal of Mathematical Science Education

EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India*.

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain*.
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK.*
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India*.
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India.*
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India*.
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India*.
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India*.
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia*.
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

Copyright Ó2009 *Technomthematics Research Foundation.*

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.
Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.
Sutra: International Journal of Mathematical Science Education (ISSN 0974 – 3340) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education

CONTENTS (Vol. 2, No. 1, 2009)

Editorial *R. Akerkar* iii

1. Tomography of Bounded Convex Domains 1 - 12
   V. K. Ohanyan and N. G. Aharonyan
2. Deceptive Patterns in Mathematics 13 - 21
   V. G. Tikekar.
3. Harnessing Geometric Manipulatives as a Revitalization Strategy for Mathematics Education in Nigeria 22 - 28
   Felix Ehizibue Aburime
4. Mathematical Puzzle. 29
5. Solutions to Earlier Puzzles. 30
6. Interesting Facts about *e.* 31
7. Useful Mathematical Symbols. 32

Sutra: International Journal of Mathematical Science Education

EDITORIAL (Vol. 2, No. 1, 2009)

I am happy to present a new issue of the Sutra journal. This issue contains three papers and other regular sections. First expository paper reviews results in tomography of bounded convex domains. In second paper you will envisage mathematical patterns and their deceptive nature. I am sure that this article will guide our reader to strengthen the understanding of the term 'sequence'. Next paper discusses mathematics performance in Nigerian secondary schools and suggests remedial solution.

Last year, we published only one issue, because we launched our journal in the month of June 2008. Now onwards, we will bring out two issues per year. In this connection, we welcome contributions, comments ad suggestions from our readers.

R. Akerkar

Editorial Coordinator

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 1,1-12, 2009

TOMOGRAPHY OF BOUNDED CONVEX DOMAINS

V. K. Ohanyan and N. G. Aharonyan
Yerevan State University, Armenia
victo@aua.am

The present paper is a review of results in tomography of bounded convex domains. In the last century German mathematician W. Blaschke formulated the problem of investigation of bounded convex domains in the plane using probabilistic methods. In particular, the problem of recognition bounded convex domains $\mathbb{D}$ by chord length distribution function (or density function). Till recently explicit expressions for the chord length distribution functions have been known in the case when $\mathbb{D}$ is a disc, a regular triangle (see R. Sulanke, 1961) and a rectangle (see W. Gille, 1988). Using delta-formalism in Pleijel identities we have obtained explicit expression of the chord length distribution function for any regular polygon (see V. K. Ohanyan and H. Harutyunyan, 2009). In the last years the notion of orientation-dependent chord length distribution function have been introduced. In the paper of N. Aharonyan (2008) an explicit formula for orientation-dependent chord length distribution function for any bounded convex domains $\mathbb{D}$ have been obtained. These questions are connected with Covariogram Problem: Does the covariogram determine a convex domain, among all convex domains, up to translations and reflections? G. Matheron in 1986 conjectured a positive answer for this problem. In fact, the covariogram problem is equivalent to the problem of determining a convex domain from all its orientation-dependent chord length distributions (see W. Nagel, 1993; G. Averkov and G. Bianchi, 2007). All these problems are the problems of geometric tomography (see R. Gardner, 1995), since orientation-dependent chord length distribution function is the probability that parallel X-ray in direction ϕ less than or equal to y.

1. Description of results

Complicated geometrical patterns occur in many areas of science. Their analysis requires creation of mathematical models and development of special mathematical tools.

The corresponding area of mathematical research is called Stochastic Geometry (see [1], [2] and [10]). Among more popular applications are Stereology and Tomography (see [14], [3], [16], [21]).

The methods of form analysis are based on analysis of the objects as figures, i.e. as subsets of the plane. For these sets, geometrical characteristics are considered that are independent of the position and orientation of the figures (hence they coincide for congruent figures). Classical examples are area and perimeter of a figure. In the late thirties of the last century German mathematician W. Blaschke formulated the problem of investigation of bounded convex domains in the plane using probabilistic methods. In particular, the problem of recognition bounded convex domains $\mathbb{D}$ by chord length distribution function. Random lines generate chords of random length in convex domain $\mathbb{D}$. The corresponding distribution (or density) function is called the chord length distribution function that we denote by $F(y)$ (or chord length density function $f(y)$). In the initial stage of investigation mathematicians tried to find explicit expressions of the chord length distribution (or density) functions for concrete domains $\mathbb{D}$ in the terms of elementary functions. Till recently explicit expressions for the chord length distribution functions have been known in the case when $\mathbb{D}$ is a disc, a regular triangle (see [6]) and a rectangle (see [7]). These results have been obtained using the definition of chord length distribution function for a domain $\mathbb{D}$ (i.e. definition of the geometric probability in the space of lines in the plane). A family of identities primarily associated with isoperimetric inequalities for planar convex domains was discovered by A. Pleijel (see [4] and [5]) in 1956. Using combinatorial principles in integral geometry (in the meaning of W. Blaschke) R. V. Ambartzumian proved these identities (see [1], [2]) and pointed out (see [2], section 6.10) that classical Pleijel identities can be used for the finding chord length distribution function for the polygons which have no parallel sides. In the recent years our group has obtained important results in this direction. Using delta-formalism in Pleijel identities we have obtained explicit expression for the chord length distribution function of a regular polygon (see [22]). In the particular

cases of a regular triangle, a square, a regular pentagon and a regular hexagon our formula for the chord length distribution function coincides with formulas available in the literature (see [6], [7], [8] and [9]) for $n = 3$, 4, 5 and 6 correspondingly. The form of the length density function is related to certain features of the corresponding figures. For example, poles of this function are related to parallel pieces of the contour and the form of $f(y)$ for y close to its maximum is essentially related to smaller details of the contour. The question is related to some methods of pattern recognition (see [17], [18] and [23]). The authors of [19] have shown that chord length distribution function does not characterize the convex domain $\mathbb{D}$. In [19] an example of two noncongruent convex domains with the same distribution function of the chord length have been constructed. The determination of the chord length distribution function has a long tradition of application to collections of bounded convex bodies forming structures in metal and ceramics. The series of formulae for chord length distribution functions may be of use in finding suitable models when empirical distribution functions are given (see [10], page 116). The results concerning certain infinite cylinders with regular polygonal bases are obtained in [11] and [15].

In the last years have been introduced the notion of orientation-dependent chord length distribution function $F_\phi(y)$, while F(y) is called mixed orientation distribution function. Using delta-formalism we have obtained orientation-dependent chord length distribution function (disintegration of mixed oriented chord length distribution function) (see [12]). These questions are connected with Covariogram Problem: Does the covariogram determine a convex domain, among all convex domains, up to translations and reflections? G. Matheron in 1986 conjectured a positive answer for this problem (see [24]). In fact, the covariogram problem is equivalent to the problem of determining a convex domain from all its orientation-dependent chord length distributions (see [13] and [20]).

At the Conference on Tomography at Oberwolfach in 1990, R. Gardner introduced the term geometric tomography. In the R. Gardner book [21], the following definition is

offered: "Geometric tomography is the area of mathematics dealing with the retrieval of information about a geometric object from data about its sections, or projections, or both". The word projection is used in the sense of a shadow, that is, the usual orthogonal projection on a line. The parallel X-ray of $\mathbb{D}$ in the direction ϕ gives for each $x \in \phi^{\perp}$ the length of the chord of intersection of $\mathbb{D}$ with the line through x parallel to ϕ. A set of four directions whose slopes have a transcendental cross ratio will ensure that the corresponding parallel X-rays determine each planar convex domain (see [21]).

2 Chord length distribution functions

Let $\mathbb{G}$ be the space of all lines g in the Euclidean plane $\mathbb{R}^2$, (p, φ) = the polar coordinates of the foot of the perpendicular to g from the origin O, be standard coordinates for a line $g \in \mathbb{G}$. Denote by $\mathbb{D}$ a convex bounded polygon in the plane and by a_1, a_2,..., a_n sides of $\mathbb{D}$. Then

$$[\mathbb{D}] = \{g \in \mathbb{G} : \ g \cap \mathbb{D} \neq \emptyset\} = \bigcup_{i<j} ([a_i] \cap [a_j])$$

where $[a_i] \cap [a_j]$ is the set of lines hitting both sides a_i and a_j of $\mathbb{D}$.

Let b_{ij} be the distance between the parallel segments a_i and a_j (i.e. the distance between the lines containing a_i and a_j). Further, $h(\varphi) \neq b_{ij}$ is the height of the maximal parallelogram with two sides equal to $\chi(\varphi) = g(\varphi) \cap \mathbb{D}$, $g(\varphi) \in [a_i] \cap [a_j]$ ($g(\varphi)$ is a line with φ-direction), and the other two sides lie on the parallel sides a_i and a_j,

$$h(\varphi_\chi) = h\left(\arccos \frac{b_{ij}}{|\chi(\varphi)|}\right) + h\left(2\,\pi - \arccos \frac{b_{ij}}{|\chi(\varphi)|}\right).$$

Hence , $h(\cdot) = 0$ if the parallelogram is empty.
For the value of φ such that $|\chi(\varphi)| = y$ we have $h(\varphi_y) = h(\varphi_\chi)$.

We obtain

$$F(y) = 1 - \frac{1}{\sum\limits_{i=1}^{n} |a_i|} \left[\sum_{i<j}^{I} \frac{y}{\sin\gamma_{ij}} \int_{\Phi_{ij}(y)} \sin\varphi \, \sin(\gamma_{ij} - \varphi) \, d\varphi - \right.$$

$$\left. - \sum_{i<j}^{II} I_{ij}(y) \, h(\varphi_y) \, \frac{\sqrt{y^2 - b_{ij}^2}}{b_{ij}} + \sum_{i=1}^{n} (|a_i| - y)^+ \right], \qquad (2.1)$$

where $\sum^I$ is over all nonparallel pairs of segments a_i, $a_j \subset \partial\mathbb{D}$ and $\sum^{II}$ is over all parallel pairs of segments a_i and $a_j \subset \partial\mathbb{D}$, while $|a_i|$ is the length of a_i, $i = 1, ..., n$.

In the case where $\partial\mathbb{D}$ contains no pairs of parallel sides, formula (2.1) coincides with the expression given by R. V. Ambartzumian in [2], page 158.

It follows from (2.1) that to find distribution function $F(y)$ we have to calculate integrals of the form

$$\frac{1}{\sin\gamma} \int_{\Phi(y)} \sin\varphi \, \sin(\gamma - \varphi) \, d\varphi$$

for any two nonparallel segments a and b ($a \leq b$) with the angle γ between a and b (or their continuations) and also calculate the second sum in (2.1) (for pairs of parallel sides). Here $\Phi(y)$ is

$$\Phi(y) = \{\varphi : \text{ a chord joining } a \text{ and } b \text{ exists with direction } \varphi \text{ and length } y\}.$$

For the proof of (2.1) see [22] and [8].

In the classical Pleijel identities integration is over the measure in the space $\mathbf{G}$ of lines which is invariant with respect to the all Euclidean motions (see [1], [2]). In the paper [12] generalized Pleijel identities for any locally-finite, bundleless measure in the space $\mathbf{G}$ have been proved. These identities are applied to find the so-called orientation-dependent chord length distribution functions for bounded convex domains:

$$b(\phi_0) \cdot [1 - F_{\phi_0}(y)] = \frac{1}{2} \int_{[\mathbb{D}]} \delta(|\chi(g)| - y) \, |\chi(g)| \, |\sin(\phi - \phi_0)| \, dg -$$

$$-\frac{1}{2} \int_{[\mathbb{D}]} \delta'(|\chi(g)| - y) \cdot |\chi(g)|^2 \, |\sin(\phi - \phi_0)| \, \cot\alpha_1 \, \cot\alpha_2 \, dg, \qquad (2.2)$$

where $b(\phi)$ is the breadth function (ϕ is a direction), $\delta(y)$ is the Dirac's δ–function concentrated at y, α_1 and α_2 are the angles between $\partial\mathbb{D}$ and the line g at the endpoints of $\chi(g) = g \cap \mathbb{D}$ which lie in one half-plane with respect to the inside of $\mathbb{D}$, while $F_\phi(y)$ is the orientation-dependent distribution function at ϕ

$$F_{\phi_0}(y) = \frac{1}{b(\phi_0)}\,\mathcal{L}_1\{p:\ |\chi(p,\phi_0)| < y\},$$

where $\mathcal{L}_1$ is one dimensional Lebesgue measure.

In the case where $\mathbb{D}$ is a polygon with no parallel sides, we have the following formula (see [12]):

$$b(\phi_0)\cdot[1 - F_{\phi_0}(y)] =$$

$$= \frac{y}{2}\sum_{i<j}\frac{1}{\sin\gamma_{ij}}\int_{\Phi_{ij}(y)}|\sin(\phi-\phi_0)|\,(\cos\phi\cos(\gamma_{ij}-\phi)+2\sin\phi\sin(\gamma_{ij}-\phi))\,d\phi+$$

$$+\sum_{i=1}^{n}|\sin(\phi_i-\phi_0)|\left[(a_i-y)^+ - \frac{a_i\,I(a_i\geq y)}{2}\right],$$

where a_i is the side of the polygon $\mathbb{D}$, $i = 1, 2, ..., n.\gamma_{ij}$ is the angle between nonparallel sides a_i and a_j (or their continuations), α_1 is the angle between a_i and $\chi(g) = g \cap D$, and α_2 is the angle between a_j and $\chi(g)$ ($g \in [a_i] \cap [a_j]$). and

$\Phi_{ij}(y) = \{\varphi:$ a chord joining a_i and a_j exists which has direction φ and its length is $y\}$.

3 KNOWN PARTICULAR CASES

I) A DISC.

For the case of the disc of radius R distribution function has the following form:

$$F_{\phi_0}(y) = \begin{cases} 0 & \text{if} \quad y \leq 0, \\ 1 - \frac{\sqrt{4R^2-y^2}}{2R} & \text{if} \quad y \in (0, 2R). \\ 1 & \text{if} \quad y > 2R \end{cases} \tag{3.1}$$

The corresponding density function is

$$f_{\phi_0}(y) = \begin{cases} 0 & \text{if} \quad y \notin (0, 2R), \\ \frac{y}{2R\sqrt{4R^2-y^2}} & \text{if} \quad y \in (0, 2R) \end{cases} \tag{3.2}$$

that coincides with the form of "mixed" density function $f(y)$ (see [10]).

II) A REGULAR TRIANGLE.

In the case of a regular triangle of side a, density function has the form:

$$f(y)=\begin{cases}0, & \text{if} \quad y\notin[0,a]\\ \frac{1}{2a}+\frac{\pi}{3\sqrt{3}a}, & \text{if} \quad 0<y\le\frac{\sqrt{3}a}{2}\\ -\frac{\sqrt{4y^2-3a^2}}{2\,y^2}+\frac{1}{2a}+\frac{2}{\sqrt{3}a}\,\arcsin\frac{a\sqrt{3}}{2y}-\frac{2\pi}{3\sqrt{3}a}, & \text{if} \quad \frac{\sqrt{3}a}{2}<y\le a\end{cases} \tag{3.3}$$

and distribution function has the form

$$F(y)=\begin{cases}0, & \text{if} \quad y\le 0\\ \left(\frac{1}{2}+\frac{\pi}{3\sqrt{3}}\right)\frac{y}{a}, & \text{if} \quad 0\le y\le\frac{\sqrt{3}a}{2}\\ \frac{y}{2a}-\frac{2\pi}{3\sqrt{3}}\frac{y}{a}+\frac{2y}{\sqrt{3}a}\,\arcsin\frac{a\sqrt{3}}{2y}+\frac{\sqrt{4y^2-3a^2}}{2y}, & \text{if} \quad \frac{a\sqrt{3}}{2}\le y\le a\\ 1, & \text{if} \quad y\ge a.\end{cases} \tag{3.4}$$

The result coincides with the result by Sulanke ([6] see also [10]).

III) A RECTANGLE.

In the case of rectangle with sides a and b, $a\le b$, density function of the length of a random chord has the form:

$$f(y)=\begin{cases}0, & \text{if} \quad y\notin[0,\sqrt{a^2+b^2})\\ \frac{1}{a+b}, & \text{if} \quad 0<y\le a\\ \frac{a^2\,b}{(a+b)y^2\sqrt{y^2-a^2}}, & \text{if} \quad a<y\le b\\ \frac{ab}{(a+b)y^2}\left(\frac{a}{\sqrt{y^2-a^2}}+\frac{b}{\sqrt{y^2-b^2}}\right)-\frac{1}{a+b}, & \text{if} \quad b<y\le\sqrt{a^2+b^2},\end{cases} \tag{3.5}$$

while distribution function $F(y)$ for a rectangle has the form

$$F(y)=\begin{cases}0, & \text{if} \quad y\le 0\\ \frac{y}{a+b}, & \text{if} \quad 0\le y\le a\\ \frac{a}{a+b}+\frac{b}{a+b}\,\frac{\sqrt{y^2-a^2}}{y}, & \text{if} \quad a\le y\le b\\ 1-\frac{y}{a+b}+\frac{1}{(a+b)y}\,(a\sqrt{y^2-b^2}+b\sqrt{y^2-a^2}), & \text{if} \quad b\le y\le\sqrt{a^2+b^2}\\ 1, & \text{if} \quad y\ge\sqrt{a^2+b^2}.\end{cases} \tag{3.6}$$

The result coincides with the result by W. Gille ([7] see also [10]).

IV) A REGULAR PENTAGON.

The chord length distribution function for a regular pentagon has the form:

$$F(y)=\begin{cases}0, & \text{if} \quad y\le 0\\ \frac{y}{a}\left(\frac{1}{2}-\frac{\pi(\sqrt{5}-1)}{5\sqrt{10+2\sqrt{5}}}\right) & \text{if} \quad 0\le y\le a\\ 1+\frac{\pi\,(5+3\sqrt{5})\,y}{5a\sqrt{10+2\sqrt{5}}}-\frac{2\,(\sqrt{5}+1)\,y}{a\sqrt{10+2\sqrt{5}}}\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}- & \\ -\frac{\sqrt{5}+1}{2}\cos\left(\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}\right), & \text{if} \quad a\le y\le \frac{a}{2}\sqrt{5+2\sqrt{5}}\\ 1+\frac{\pi\,(5+3\sqrt{5})\,y}{5a\sqrt{10+2\sqrt{5}}}-\frac{2\,(\sqrt{5}+1)\,y}{a\sqrt{10+2\sqrt{5}}}\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}- & \\ -\frac{\sqrt{5}+1}{2}\cos\left(\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}\right)- & \\ -\frac{(\sqrt{5}+1)^2 y}{a\sqrt{10+2\sqrt{5}}}\arccos\frac{a\sqrt{10+2\sqrt{5}}}{2y(\sqrt{5}-1)}+ & \\ +\frac{2(\sqrt{5}+1)}{(\sqrt{5}-1)^2}\sin\left(\arccos\frac{a\sqrt{10+2\sqrt{5}}}{2y\,(\sqrt{5}-1)}\right), & \text{if} \quad \frac{a}{2}\sqrt{5+2\sqrt{5}}\le y\le\frac{\sqrt{5}+1}{2}\,a\\ 1, & \text{if} \quad y\ge\frac{\sqrt{5}+1}{2}\,a,\end{cases} \tag{3.7}$$

where a is the side of a regular pentagon.

It is not difficult to verify that this function is continuous function.

The density function has the form:

$$f(y)=\begin{cases}0, & \text{if} \quad y\notin\left[0,\frac{\sqrt{5}+1}{2}\,a\right]\\ \frac{1}{a}\left(\frac{1}{2}-\frac{\pi(\sqrt{5}-1)}{5\sqrt{10+2\sqrt{5}}}\right), & \text{if} \quad 0<y\le a\\ \frac{\pi\,(5+3\sqrt{5})}{5a\sqrt{10+2\sqrt{5}}}-\frac{2\,(\sqrt{5}+1)}{a\sqrt{10+2\sqrt{5}}}\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}+ & \\ +\frac{2\,(\sqrt{5}+1)}{\sqrt{16y^2-a^2(10+2\sqrt{5})}}-\frac{a^2\,(5+3\sqrt{5})}{2\,y^2\sqrt{16\,y^2-a^2(10+2\sqrt{5})}}, & \text{if} \quad a<y\le\frac{a}{2}\sqrt{5+2\sqrt{5}}\\ \frac{\pi\,(5+3\sqrt{5})}{5a\sqrt{10+2\sqrt{5}}}-\frac{2\,(\sqrt{5}+1)}{a\sqrt{10+2\sqrt{5}}}\arcsin\frac{a\sqrt{10+2\sqrt{5}}}{4y}+ & \\ +\frac{2\,(\sqrt{5}+1)}{\sqrt{16y^2-a^2(10+2\sqrt{5}}}-\frac{a^2\,(5+3\sqrt{5})}{2\,y^2\sqrt{16\,y^2-a^2(10+2\sqrt{5})}}- & \\ -\frac{(\sqrt{5}+1)^2}{a\sqrt{10+2\sqrt{5}}}\arccos\frac{a\sqrt{10+2\sqrt{5}}}{2y(\sqrt{5}-1)}- & \\ -\frac{(\sqrt{5}+1)^2}{\sqrt{4\,y^2(\sqrt{5}-1)^2-a^2(10+2\sqrt{5})}}+ & \\ +\frac{4\,a^2(5+3\sqrt{5})}{y^2\,(\sqrt{5}-1)^3\sqrt{4\,y^2(\sqrt{5}-1)^2-a^2(10+2\sqrt{5})}}, & \text{if} \quad \frac{a}{2}\sqrt{5+2\sqrt{5}}<y\le\frac{\sqrt{5}+1}{2}\,a\end{cases} \tag{3.8}$$

The result coincides with the result by V.K. Ohanyan and N. G. Aharonyan ([8]).

V) A REGULAR HEXAGON.

In the paper [9] H. S. Harutyunyan calculated explicit expression for chord length distribution function for a regular hexagon:

$$F(y)=\begin{cases}0, & \text{if} \quad y\le 0,\\ \frac{y}{a}\left(\frac{1}{2}-\frac{\pi}{6\sqrt{3}}\right), & \text{if} \quad 0<y\le a,\\ 1+\frac{\pi\, y}{2a\sqrt{3}}-\frac{2\,y}{a\sqrt{3}}\arcsin\frac{a\sqrt{3}}{2\,y}-\frac{\sqrt{4y^2-3a^2}}{2y}, & \text{if} \quad a<y\le a\sqrt{3},\\ 1+\frac{y}{a}\left(\frac{\pi}{6\sqrt{3}}-\frac{1}{2}\right)-\frac{y}{a\sqrt{3}}\arccos\frac{a\sqrt{3}}{y}+\frac{2\sqrt{y^2-3a^2}}{y}, & \text{if} \quad a\sqrt{3}<y\le 2\,a,\\ 1, & \text{if} \quad y>2\,a,\end{cases} \tag{3.9}$$

where a is the side of a regular hexagon.

The density function has the form:

$$f(y)=\begin{cases}0, & \text{if} \quad y\notin(o,2\,\pi],\\ \frac{1}{a}\left(\frac{1}{2}-\frac{\pi}{6\sqrt{3}}\right), & \text{if} \quad 0<y\le a,\\ \frac{\pi}{2a\sqrt{3}}-\frac{2}{a\sqrt{3}}\arcsin\frac{a\sqrt{3}}{2\,y}+\frac{\sqrt{4y^2-3a^2}}{2y^2}, & \text{if} \quad a<y\le a\sqrt{3},\\ \frac{1}{a}\left(\frac{\pi}{6\sqrt{3}}-\frac{1}{2}\right)-\frac{1}{a\sqrt{3}}\arccos\frac{a\sqrt{3}}{y}+\frac{6\,a^2-y^2}{y^2\sqrt{y^2-3a^2}}, & \text{if} \quad a\sqrt{3}<y\le 2\,a.\end{cases} \tag{3.10}$$

4 Conclusion

The chord length distribution function $F(y)$ are independent of the positions of the domains in the plane, thus it coincides for congruent domains.

Assume now to have the information about the distribution of the chord length not in the "completely mixed" form, but separated direction by direction. One can show that the problem of determining a body from this data, gave a positive answer using "orientation dependent" chord length distribution, when $\mathbb{D}$ is a planar convex polygon (see W. Nagel [13]) and for any bounded convex domain (see [20]). In fact, the covariogram problem is equivalent to the problem of determining a convex body from all its separate chord length distributions (see [13], [20]). A set of four directions whose slopes have a transcendental cross ratio will ensure that the corresponding parallel X-rays determine each planar convex domain (see [21]). The following problem arise.

OPEN PROBLEM. Does there exist a finite set of directions $V = \{\phi_1, ..., \phi_m\}$ such that the corresponding set of "orientation-dependent" chord length distribution functions $F_{\phi_1}(y)$, ..., $F_{\phi_n}(y)$ determine a bounded convex domain uniquely. Find minimal m which satisfies this condition.

References

[1] Ambartzumian, R. V. (1982). *Combinatorial Integral Geometry with Applications to Mathematical Stereology*, John Wiley and Sons, Chichester.

[2] Ambartzumian, R. V. (1990). *Factorization Calculus and Geometric Probability*, Cambridge University Press, Cambridge.

[3] Ohanyan, V. K. (2008). Combinatorial principles in Stochastic Geometry: A Review. In *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)* **43**, no. 1, 44-60, (Distributed by Springer, see www.springerlink.com).

[4] Pleijel, A. (1956). Zwei kurze Beweise der isoperimetrischen Ungleichung, In *Archiv Math.* **7**, 317-319.

[5] Pleijel, A. (1956). Zwei kennzeichende Kreiseigenschaften der Kreis, In *Archiv Math.* **7**, 420-424.

[6] Sulanke, R. (1961). Die Verteilung der Sehnenlängen an ebenen und räumlichen Figuren. In *Math. Nachr.* **23,** 51–74.

[7] Gille, W. (1988). The chord length distribution of parallelepipeds with their limiting cases. In *Exp. Techn. Phys.* **36,** 197–208.

[8] Aharonyan, N. G. and Ohanyan, V. K. (2005). Chord length distribution functions for polygons. In *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences).* **40,** (no. 4), 43–56.

[9] Harutyunyan, H. S. (2007). Chord length distribution function for regular hexagon, In *Uchenie Zapiski, Yerevan State University*, **1**, 17–24.

[10] Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*, John Wiley& Sons, Chichester, New York, Brisbane, Toronto, Singapore.

[11] Sukiasian, H. S. and Gille, W. (2007). Relation between the chord length distribution of an infinitely long cylinder and that of its base. In *J. Math. Physics* **48** 053305 1–8.

[12] Aharonyan, N. G. (2008). Generalized Pleijel Identity. In *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)* **43**, no. 5, 3-12.

[13] Nagel, W. (1993). Orientation-dependent chord length distributions characterize convex polygons, In *J. Appl. Prob.* **30**, 730 – 736.

[14] Ambartzumian, R. V. (1999). Analytical implications of combinatorial integral geometry: A Review. In *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)*, **34**, no. 6, 2 – 46.

[15] Gille, W., Aharonyan, N. G. and Harutyunyan, H. S.(2009). SAS structure functions of pentagonal and hexagonal rods. To appear in *Journal of Applied Cristallography*, vol. 42.

[16] Ambartzumian, R. V. (2009). Tomography models of random convex polygons. In *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)*, **44**, no. 1, 25 – 35.

[17] Gates, J. (1982). Recognition of triangles and quadrilaterals by chord length distribution.In *J. Appl. Prob.* **19,** 873–879.

[18] Gates, J. (1987). Some properties of chord length distributions. In *J. Appl. Prob.* **24**, 863–874.

[19] Mallows, C. and Clark, J.(1970). Linear-intercept distributions do not characterize plane sets. *J. Appl. Prob.* **7**, 240–244 (Correction (1971). **8**, 208–209).

[20] Averkov, G. and Bianchi, G. (2007). Retrieving convex bodies from restricted covariogram functions. *Adv. Appl. Prob. (SGSA)* **39,** 613–629.

[21] Gardner, R. J. (1995). *Geometric Tomography*, Cambridge University Press, New York.

[22] Ohanyan, V. K. and Harutyunyan, H.S. (2009). Chord length distribution functions for regular polygons, To appear In *Advances in Applied Probability.*

[23] Novikoff, A. B. J. (1962). Integral geometry as a tool in pattern perception, In *Principles of Self-Organization* (H. von Foerster and G. W. Zopf, Jr., Eds), Pergamon Press, London, 347-368.

[24] Matheron, G. (1986). Le covariogramme géometrique des compacts convexes de $\mathbb{R}^2$, Technical report 2/86. In *Centre de Géostatistique Ecole des Mines de Paris*, Paris **54**.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 1, 13 – 21, 2009

DECEPTIVE PATTERNS IN MATHEMATICS

[2nd title: Deceptive and Misunderstanding Nature of Patterns in Mathematics]

V. G. Tikekar*
vgtikekar@gmail.com

Mathematics is sometimes described as a study of patterns. Patterns in mathematics can be deceptive and therefore can lead to incorrect conclusion regarding the general term. Similarly patterns may not reveal correctly the nature of their members. This comes in the way of proper understanding of the mathematical concepts, like sequence. With some concrete examples we throw, in this expository article, searching light on the patterns and their deceptive nature and help the reader to consolidate the correct understanding and appreciation of the term 'sequence'.

Keywords: sequence, generating-function, primes, circles, Venn-diagrams.

1. INTRODUCTION

Nature is full of patterns. Some patterns are pleasing, while some others are challenging. Some people define mathematics as 'study of patterns.' By observing part of a pattern, can we recognize or identify the remaining part of that pattern? Many patterns deceive us, and we get surprised by the unexpected turn that the pattern takes in the part that follows the part of that pattern that was seen. Especially in mathematics it is all the more challenging to find the general form, of any pattern, that will unfold or generate all the observed individual members of the pattern. In this article, we shall study and examine some patterns of such type.

2. NEXT MEMBER IN A PATTERN

Consider the pattern whose initial terms are

$$2, 4, 6, 8,\ldots\ldots \qquad -(1)$$

What is the next term in this pattern? This type of question in which some initial terms are given and we are asked to recognize the next term, is very common in many examinations meant not only for mathematics students, but also in examinations designed to select even clerks. Tests in which such questions are set are many times called 'intelligence tests', or 'psychology tests', or 'common-sense tests'. Readers will draw their own conclusion, after reading this article, as to how far these adjectives used for such tests are right.

What is the answer to the question which asks the next term in (1)? Most people would say 10. To see whether this answer is correct or not, we must examine the reasoning that would be put forward by those who give the answer 10. Different persons giving the same answer 10,

*C/O Dr Thakar, New Mahadwar Road, Near Padmaraje School, Kolhapur – 416012 (India)

could give the reasoning in different ways. Some would say, "In (1) we record the number 2, drop the next number 3, then record the next number 4, again ignore the next number 5, record 6,omit 7, and record 8. So we must follow this pattern, discard the next number 9, and record 10. Hence the answer is 10." Some others would say, "Starting with 2, we are noting the alternate integers. Therefore the answer is 10." Somebody else may argue, "We are writing even numbers starting with 2. The even number after 8 is 10. So the required answer is 10." All these arguments convey the same reasoning even though they are worded differently. Some other mathematically inclined person may state the same reasoning as follows: "Numbers given in (1) are generated by the formula 2n. If we substitute
n = 1,2,3,4 in 2n we get 2, 4, 6, 8 respectively. So if we put the next value of n, namely 5, in 2n, we get 10. Hence 10 is the expected answer. Thus 2n can be taken as the generating function of the pattern given by (1)."

Now consider the formula

$$n^4 - 10n^3 + 35n^2 - 48n + 24 \quad -- \quad (2)$$

If we put n = 1,2,3,4 in (2) we get 2,4,6,8 respectively. So why not consider (2) as the generating function of (1). In that case, if we put n = 5 in (2) we get 34. So 34 can be taken as the answer to the query "what is the next term in (1)?"

One may be surprised with this and may wonder how one discovered the formula (2). It is easy to see that we want a formula (or expression) in n that will generate the same numbers (viz., 2,4,6,8) that are generated by the formula 2n, when we substitute n = 1,2,3,4 in it. So to get such a formula we must add to (or subtract from) 2n the term(s) that would become zero when n = 1,2,3,4. So

$$2n + (n-1)(n-2)(n-3)(n-4) \quad -- (3)$$

will be that formula. If we simplify (3) we get (2). Not only this, but we can construct infinite number of such formulae, that will generate 2,4,6,8 when we put n = 1,2,3,4 respectively in each one of them. To see this, look at

$$2n + a\,(n-1)(n-2)(n-3)(n-4) \quad -- (4)$$

where a can be any one of the values 2, 3, 4, 5, ……. (in fact, any number for that matter), and it is easy to see that in (4) we have got that infinite number of formulae, since for any such value of a we get 2, 4, 6, 8 when we put n = 1, 2, 3, 4 respectively in (4).

The moral of all the discussion above is that there cannot be a unique answer to the question posed at the beginning of this section, and so, such question should not be asked, really speaking. Many, not so enlightened by the mathematical precision of thinking, would call (1) a sequence. In the study of Mathematics 'sequence' is a very important and basic concept, and what is very important to know, understand, and remember (as has been very clear from our example under discussion) is that

A SEQUENCE CANNOT BE DEFINED BY ENUMERATING FINITE NUMBER OF TERMS EVEN THOUGH THEY ARE LARGE IN NUMBER.
In other words, SEQUENCE DOES NOT GET DEFINED UNLESS ITS GENERAL TERM IS STATED. --- (5)

In fact, the fact (5) is very much embodied in the definition of the term 'sequence' which states that 'sequence is a function whose domain is the set of natural numbers.' But it is rarely properly understood, as is clear from the example we have discussed just now. Thus

2, 4, 6, 8,…….. , 2n, …….. --- (6)

is a sequence. Similarly,

2, 4, 6, 8, …… , (n^4 – 10n^3 + 35n^2 – 48n + 24), …… -- (7)

is also a sequence which is different from the one in (6). But (1) is not a sequence.

One can easily see that even if we enumerate large number of terms, finite in number, in (1), we can always construct many general formulae, like the ones in (3) or (4) each of which will generate all those large number of terms.

Thus mere enumeration of a finite number of terms (without giving the general term), one followed by another, forming a pattern could deceive us, as is done by (1), if we try to find a general term that generates all those given finite number of terms. One has to be careful in such tasks.

One can see a simile here. We have, on our earth, billions of human beings, each having one nose, one mouth, two ears, and two eyes; but each one is different, and can be identified, distinguished, and recognized as different. Similarly we have billions of sequences each having its first four terms as 2, 4, 6, 8. In order to recognize, distinguish, and identify each one of them, we have to specify the general term like 2n, or n^4 – 10n^3 + 35n^2 – 48n + 24, or the like. In fact, unless such a general term is specified, it is not eligible to be called or named as a sequence in Mathematics. We have not to ignore and forget this lesson brought forward by all the discussion of this section.

3. FERMAT GETS DECEIVED

Many great mathematicians have also been deceived, by enumeration of finite number of terms, in recognizing and identifying the pattern exhibited by them. For example, Fermat (1601 – 1665), a great French Mathematician saw the following five numbers in that order:

3, 5, 17, 257, 65537 -- (8)

Fermat noticed that these are prime numbers and can be generated by the formula

{2^(2^n)} + 1 -- (9)

when we substitute n = 0, 1, 2, 3, and 4 respectively, in it. So Fermat contemplated and thought for a while and declared that (9) would serve as a generating function for primes. He must have been delighted to feel that he discovered, by examining the pattern exhibited by (8), a formula that generates primes, an important achievement in the study of primes. But alas! His examination of (8) deceived him, This became evident when another great mathematician Euler (1707 – 1783) pointed out, in 1737 [i.e. about 100 years after Fermat's examination of the pattern (8)] that if we put n = 5 in (9), we get

{2^(2^5)} + 1 = 641 x 6700417 -- (10)

which, therefore, is not a prime. In fact, no integer n (> 5) has yet been found such that {2^(2^n)} + 1 is a prime. So one learns an exceedingly important lesson that a pattern displayed by enumeration of finite number of terms can turn out to be deceptive in conveying its nature and hence it is absolutely necessary to be extra careful, cautious, and vigilant while examining patterns in mathematics.

4. MISTAKEN JUDGEMENT OF POLIGNAC

We shall cite one more example of deceit, of one more mathematician, by the pattern exhibited by the following enumeration of the finite number of terms:

3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57 --- (11)

Polignac, a French mathematician, in the year near about 1850, observed the following pattern in (11):

3 = 2^0 + 2; 5 = 2^1 + 3; 7 = 2^1 + 5 = 2^2 + 3; 9 = 2^1 + 7 = 2^2 + 5; 11 = 2^2 +7 = 2^3 + 3; 13 = 2^1 + 11 = 2^3 + 5; 15 = 2^1 + 13 = 2^2 +11 = 2^3 + 7; 17 = 2^2 + 13; 19 = 2^3 + 11 = 2^4 + 3;

21 = 2^1 + 19 = 2^2 + 17 = 2^3 + 13 = 2^4 + 5; 23 = 2^2 + 19 = 2^4 + 7; etc. etc. --- (12)

From this observation, he concluded that

EVERY ODD POSITIVE INTEGER CAN BE EXPRESSED AS A SUM A PERFECT POWER OF TWO AND A PRIME NUMBER. --- (13)

Here again Polignac was deceived in his process of observation, though he had actually claimed that he had examined the odd integers up to ten million. Later a critical eye of another investigator noticed that the property exhibited in (12) is not possessed by every odd positive integer, as claimed by Polignac, since he could discover a counterexample, namely 127. We can convince ourselves about this fact by noticing that

$$127 = 2^0 + (2 \times 63) = 2^1 + (5 \times 25) = 2^2 + (3 \times 41) = 2^3 + (7 \times 17)$$
$$= 2^4 + (3 \times 37) = 2^5 + (5 \times 19) = 2^6 + (7 \times 9) \quad \text{--- (14)}$$

and noting that $2^7 = 128 > 127$ so that only powers of 2 that need be taken into consideration for this verification are 0 through 7.

Thus Polignac was not cautious and attentive enough in examining (11) before formulating and consolidating his conclusion, or in other words, one can say that patterns in (11) and (12) deceived him to come to a hasty conclusion (13).

5. PARTITION OF CIRCLE'S INTERIOR

Let us now take an illustration from Geometry. Consider a circle. Take n points on its circumference. Join every pair of points so chosen by a line segment. Each such line segment will be a chord of that circle. The points on the circle should be so taken that the chords when drawn as described above would divide the interior of the circle in maximum possible number of compartments (i.e. parts). Count the maximum number of parts into which the circle can be partitioned by such a construction. Look at the Figures 1(a) through 1(e) in which n = 1, 2, 3, 4, 5 respectively. The maximum number of parts in which a circle can be divided by such construction is noted in Table 1.

TABLE 1:

n	1	2	3	4	5	
Maxi number of parts into which circle is partitioned	1	2	4	8	16	-- (15)

So what would be the maximum number of parts into which the interior of a circle will be partitioned if we take n = 6 points on it and carry out the construction described above? Members of the pattern exhibited in (15) can be generated by the formula

$$2^{(n-1)} \quad \text{-- (16)}$$

So one would like to conclude that maximum number of compartments in which a circle is partitioned by the chords drawn by joining every pair of points selected from the n points situated on the circle is given by the formula (16) and hence the number of such compartments will be $2^{(6-1)} = 2^5 = 32$ when we take 6 points on a circle and draw the chords as stated above.

But if we actually draw such a figure, as is done in Figure 2, we realize that one cannot get more than 31 compartments when n = 6. In fact, (16) is not the correct formula in this situation and we have been deceived, in our examination, by the pattern displayed in (15). One can prove mathematically that the members in the pattern (15) can also be generated by the formula

$$(1/24)(n^4 - 6n^3 + 23\,n^2 - 18n + 24) \quad \text{-- (17)}$$

and that (17) really gives the correct formula for the problem on hand here. Hence the maximum number of compartments in which a circle can be partitioned is 31 when 6 points are taken on the circle and the construction of drawing chords is carried out as described above. Is this not a glaring case, of deception by a pattern, warning us to be extremely cautious and careful in dealing with patterns, however simple they may look?

6. CIRCLES AND VENN-DIAGRAMS

We shall study one more concrete illustration related to drawing of figures. Venn diagrams are employed to pictorially represent sets. We generally use different circles to represent different individual sets while drawing Venn diagrams. See Figures 3, 4, 5 in which 1, 2, and 3 sets (named using letters A, B, C) are shown respectively (using circles) as subsets of a universal set U represented by a rectangle following usual convention and practice. We notice that

UNIVERSAL SET IS PARTITIONED, AS WOULD BE EXPECTED, INTO $2^1 = 2$, $2^2 = 4$, and $2^3 = 8$ PARTS (SUBSETS) IN Figures 3, 4, and 5 RESPECTIVELY BY THE CIRCLES IN THEM and these numbers display a certain pattern. -- (18)

When we have to deal with 4 sets, we know that the maximum number of subsets into which the universal set should get partitioned is $16 = 2^4$. From the pattern observed in Figures 3, 4, 5 we expect that when 4 different circles would represent 4 different sets intersecting each other, the maximum number of compartments into which the universal set would get partitioned must be $2^4 = 16$. But again this turns out to be an eye-opening case of deception carried out by the patterns displayed by the Figures 3, 4, 5. Look at the Figure 6, and learn from it the fact

YOU MAY DRAW 4 INTERSECTING CIRCLES, REPRESENTING 4 SETS, ANY WAY YOU PLEASE, BUT IN THAT CASE THE MAXIMUM NUMBER OF COMPARTMENTS INTO WHICH THE UNIVERSAL SET WOULD BE PARTITIONED CANNOT EXCEED 14. – (19)

This deception teaches us an important lesson, namely,

WHEN WE WANT TO DEAL WITH MORE THAN 3 SETS, CIRCLES ALONE SHOULD NOT BE EMPLOYED TO REPRESENT THE VARIOUS SETS WHILE DRAWING VENN DIAGRAMS. – (20)

Otherwise the practical situation represented by 4 or more sets will not get correctly reflected in pictorial representation carried out by Venn diagrams. Thus

WHEN WE DEAL WITH 4 OR MORE SETS, WE ARE PER-FORCE REQUIRED TO EMPLOY DIFFERENT FIGURES LIKE RECTANGLES, OVALS, CIRCLES, ETC. TO REPRESENT DIFFERENT SETS WHILE USING VENN-DIGRAM REPRESENTATION in order to get real situation correctly and properly reflected in the pictorial representation. — (21)

7. CONCLUSION

Thus, in conclusion, we state that though Mathematics, in some sense, can be considered as a study of patterns, one must be aware of deceptive nature of patterns and should be extremely cautious and careful in studying, observing, distinguishing, and recognizing patterns, and especially in identifying the general term from the given or known finite number of terms of any pattern that is observed or noticed by us.. It is important that we must not accept any formula as a general term of a given pattern representing a certain practical situation unless it is mathematically proved to be so.

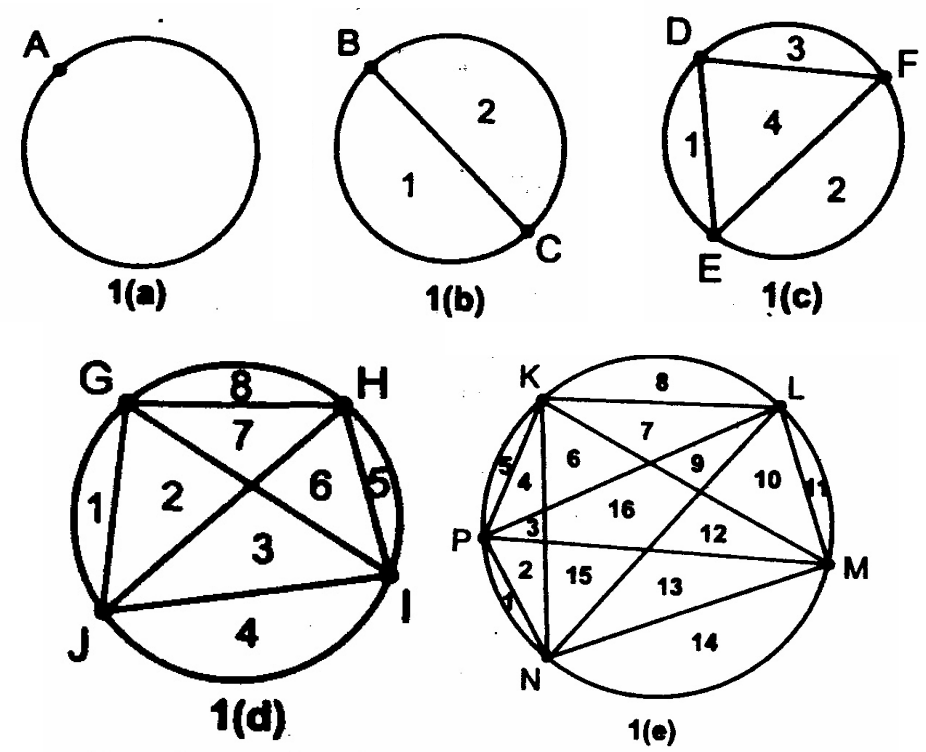


Figure 1

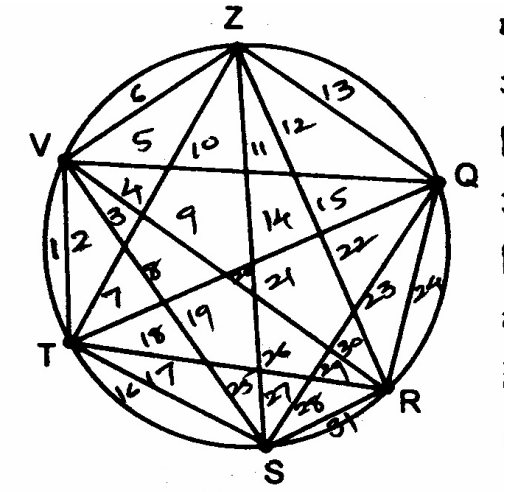


Figure 2

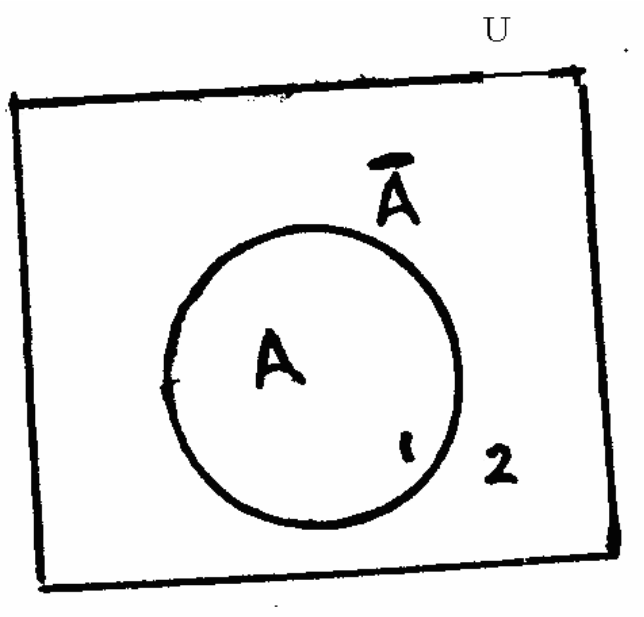


Figure 3

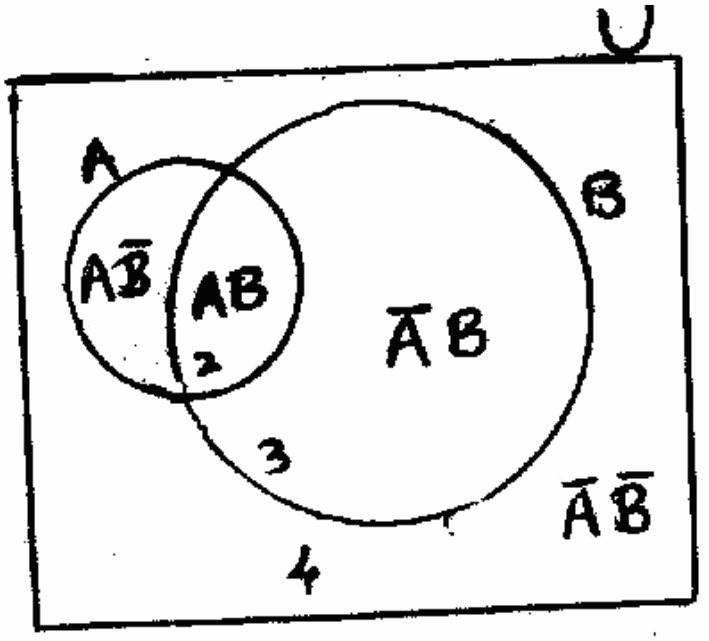


Figure 4

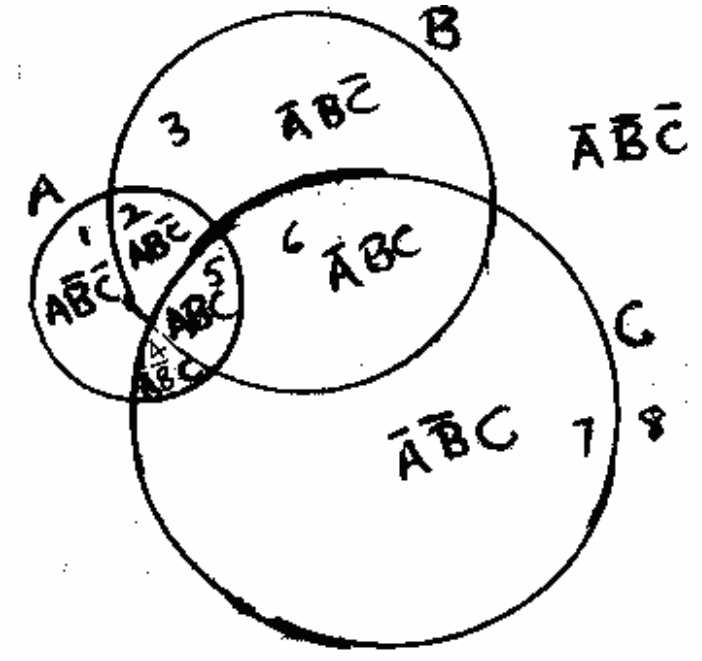


Figure 5

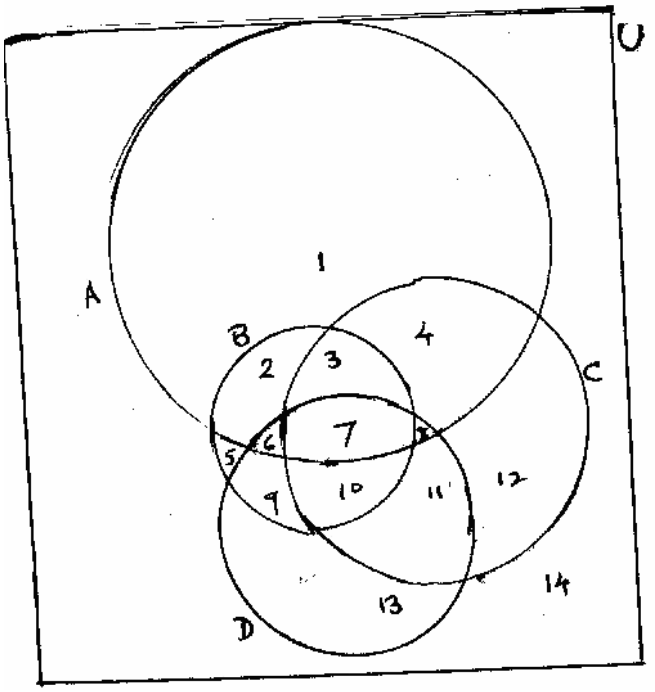


Figure 6

Bibliography

1. John H. Conway and Richard K. Guy: The Book of Numbers, Copernicus (An imprint of Springer-verlag), 1996.
2. Jeff Young and Duncan A. Buell: The twentieth Fermat number is composite, Math. Comput. 50 (1988), 261-263.
3. Richard K. Guy: The Strong Law of Large Numbers, Amer. Math. Monthly 95 (1988),697-712

Brief biographical sketch of the author:

Dr V. G. Tikekar is a retired professor and chairman of Mathematics Department at Indian Institute of Science (IISc), Bangalore. He visited UK and USA under Commonwealth Academic Staff Fellowship, and Indo-American Fellowship respectively. His interests are Numerical Mathematics, Mathematical Programming, Theoretical Computer Science, Probability, and Statistics. He is a recipient of Jaya Jayant award for teaching excellence 1989 at IISc.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 1, 22-28, 2009

HARNESSING GEOMETRIC MANIPULATIVES AS A REVITALIZATION STRATEGY FOR MATHEMATICS EDUCATION IN NIGERIA

Felix Ehizibue Aburime

Department of Mathematics Education, Benson Idahosa University,
Benin City, Edo State, Nigeria
felab2000@yahoo.com

This study tested out the implication of harnessing geometric manipulatives on the mathematics achievement of secondary school students in Nigeria. An experiment was conducted in Edo State, using a multistage sample of students in three secondary schools. The instruments for the research were a testing instrument, Mathematics Achievement Test (MAT) and a treatment instrument, Simple Improvised Manipulatives (SIM). Statistical analysis of pretest and posttest scores was carried out by means of SPSS computer software. The independent samples t-test was the statistical method adopted. Positive and statistically significant results indicated that geometric manipulatives could be harnessed as a revitalization strategy for mathematics education in Nigeria.

Keywords: Mathematics Education; Geometric Manipulatives; Nigeria.

I. – Introduction

Perennial poor mathematics performance in Nigerian secondary schools has generated an overwhelming need for a review of current teaching and learning strategies. The Mathematical Association of Nigeria (MAN) once declared a War Against Poor Achievement in Mathematics (WAPAM). Unfortunately, WAPAM achieved little in reversing the trend of poor mathematics achievement in Nigerian secondary schools. Thus, this paper aims to proffer a new strategy towards reversing the trend.

To achieve its aim, the remainder of the paper is organized in the following manner. The next section builds up a theoretical framework for the study. Section 3 outlines the investigation methods. Section 4 presents the results. Section 5 concludes the paper.

II. –Theoretical Framework

1. Poor performance in mathematics in nigeria

Poor achievement in mathematics in Nigerian secondary schools has assumed alarming proportions and caused a lot of concern for many years. [Adeniji (1998)] and [Amoo (2001)] have expressed similar tales of woes about low achievement in mathematics in Nigerian secondary schools. This problem is clearly evident in the following table.

Table 1 – Performance in Mathematics in West African School Certificate Examinations (1995-1997)

Year	Entry	Candidates who sat for the exam	Credit 1 – 6	Pass 7 – 8	Fail 9
1995	466971 %	462273 99	76080 16.5	185931 40.2	200262 43.3
1996	519656 %	514342 99.07	51587 10.01	190899 37.07	272356 52.9
1997	621844 %	616923 99.2	47252 7.65	161526 26.18	408145 66.16

Source: Amoo (2001)

From the table, failure in mathematics (F9) increased from 43.3% in 1995 to 52.9% in 1996, and then increased again to 66.16% in 1997. The seriousness of the adverse implications of these results is better appreciated when one glances through the mathematics admission requirements for Nigerian universities, as shown in the next table.

Table 2 – Mathematics Admission Requirements for Nigerian Universities

S/N	Faculty	Number of courses available	Courses needing mathematics	% Needing mathematics
1	Administration	14	11	79
2	Agriculture	36	36	100
3	Arts	56	0	0
4	Education	73	42	58
5	Engineering	53	53	100
6	Law	6	0	0
7	Medical Sciences	15	15	100
8	Science	60	60	100
9	Social Sciences	27	10	37
Total		340	227	67

Source: Joint Admissions and Matriculation Board (1998)

From the table, out of 340 courses available in Nigerian universities in 1999 / 2000 academic session, 227 courses (67%) needed a credit (1-6) pass in mathematics for any student to be considered for admission.

While Table 1 shows that the percentage of students with credit pass in mathematics dropped from 16.5% in 1995 to 10.01% in 1996, and further dropped to 7.65% in 1997, Table 2 shows that, in the faculties of agriculture, engineering, medical sciences and science, all courses required credit pass in mathematics as an admission requirement. It is therefore obvious that, though mathematics is very important in the Nigerian educational system, the performance in mathematics has been very low. This problem of very low mathematics achievement could adversely affect national development,

especially in science and technology.

2. Manipulatives – a plausible panacea for poor performance in mathematics

According to [National Science Foundation (2002)], manipulatives are materials designed to provide concrete experiences that can help students make the link between mathematical concepts and the real world. They are objects that aid children in visualizing mathematical processes. However, they are also useful in adult education. In the main, improvised geometric manipulatives are made from ordinary cardboard paper. This makes them cheap and easily available. Teachers, and even students, can easily produce a good number of them locally. However, analyzing the effective application of improvised manipulatives, [Ivowi (1999: 483)] has cautioned that the practical nature of science and mathematics needs to be emphasized in order to enhance their functionality as a means of instruction.

If the use of simple improvised geometric manipulatives in this present study is found to be significantly effective in improving mathematics achievement in Nigerian secondary schools, then their widespread use could revitalize mathematics education in Nigeria and, by implication, other developing countries having similar characteristics. For any developing nation, more success in mathematics education would have the long-term effect of greater national development, especially in science and technology.

III – Investigation Methods

To analyze the possibility of harnessing geometric manipulatives as a revitalization strategy for mathematics education in Nigeria, an experiment was conducted in Edo State using a multistage sample [(Eze, 1999: 55)] of students in three secondary schools – Momodu College (MC), Agbede (mixed),
Girls Model Secondary School (GMSS), Ubiaja (girls only), and Immaculate Conception College (ICC), Benin City (boys only).

1. Research instruments

The instruments for the research were a testing instrument, Mathematics Achievement Test (MAT) and a treatment instrument, Simple Improvised Manipulatives (SIM). MAT was a multiple-choice type
of achievement test which was used for both pretest and posttest. To construct MAT, a table of specifications (or test blueprint) was drawn up for sixty-eight test items (Table 3).

Table 3 – Specifications for MAT

Mathematics Areas	**Recall of Information**	**Understanding Concepts**	**Application of Concepts**	**Total**
Geometry	5	13	8	26
Algebra	3	7	4	14
Statistics	3	7	4	14
Number and Numeration	3	7	4	14
Total	14	34	20	68

Writing of test items was followed by face and content validation, then item analysis. The face and content validation reduced the items from sixty-eight to sixty-one, while item analysis reduced the test items from sixty-one to fifty-two. The fifty-two surviving items were administered on thirty-one students in a pilot test. Kuder-Richardson formula (KR21) was applied to scores in order to measure internal consistency.

The internal consistency coefficient was 0.79; and this was considered high enough to accept MAT for research. A test-retest of students, with an interval of three weeks, yielded scores that were paired and analyzed to obtain 0.82 as test-retest reliability coefficient for MAT. Again, this was considered high enough to accept MAT as a reliable research instrument.

The treatment instrument (SIM) comprised eighteen different geometrical shapes constructed from ordinary cardboard paper. Four shapes were triangles – equilateral triangle, isosceles triangle, right-angled triangle, and scalene triangle. Six shapes were quadrilaterals – square, rhombus, rectangle, parallelogram, trapezium, and kite. There were four other plane shapes (pentagon, hexagon, circle, and semi-circle)

and four solid shapes (cube, cuboid, triangular prism, and cylinder).

2. Data collection

Data collection started with a pretest administered to students in an experimental and a control group. The pretest scores were carefully kept aside for future use. After the pretest, ten weeks of teaching followed. During this period, students in the experimental group were taught with SIM; while students in the control group were taught without SIM. Special care was taken to ensure that, in each school, the same mathematics teacher taught the experimental group and the control group. This was to eliminate what is known as the "*teacher effect*". Another teaching precaution was to ensure that whenever one group was having a mathematics class, the other group was occupied by another teacher. This was to ensure that students did not study mathematics in a class that was not theirs. After the pretest and ten weeks

of teaching, a posttest was given to all students in the experimental and control groups. The posttest scores were collated for analysis in conjunction with the pretest scores.

Table 4 shows the distribution of students in the experimental and control group in the three schools where the research took place. The experimental group comprised ninety-four students; while the control group comprised ninety-one students.

Table 4 – Distribution of Students in the Experimental and Control Group

Group	ICC	GMSS	MC	**Total**
Experimental Group	36	36	22	94
Control Group	33	35	23	91
Total	69	71	45	185

It is noteworthy that though the validated MAT was used for both pretest and posttest, this was done with three precautions. The first precaution was that MAT was administered to students in two versions. The second version was a rearrangement of the question numbers and alternative answers in the first version. Thus, question 5 with correct answer E in the first version could become question 23 with

correct answer C in the second version. The second precaution was aimed at minimizing the chances of obtaining fake scores from students who merely copy from their neighbors. In order to achieve this, question papers were given to students in a checkerboard fashion as illustrated in Table 5. In the table, '1' applies to students taking the first version of MAT, while '2' applies to students taking the second version of MAT.

Table 5 – Checkerboard Arrangement of Question Papers

1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1

Checkerboard arrangement of question papers guaranteed that each student was surrounded by other students writing a different version of MAT. The third precaution was that no student wrote the same version of MAT for pretest and posttest.

3. Data analysis

Statistical analysis of pretest and posttest scores was carried out by means of SPSS computer software. The independent samples t-test [(Warner, 2008)] was the statistical method adopted. During the data screening process, extreme outliers were identified and expunged from the data set. All the various assumptions of the independent samples t-test were satisfied.

IV. The Results

Table 6 – T-test Results for Pretest Scores

Group	**Number of Students**	**Mean**	**Standard Deviation**	**Degrees of Freedom**	**T-Calculated**	**T-Critical**	**Decision**
Experimental	94	8.74	5.37	184	0.09	1.96	Insignificant difference
Control	91	8.81	5.18				

Table 6 shows that, at the pretest level, the calculated t value (0.09) was very much below the critical t value (1.96). This also had no effect, as indexed by $?^2$ [(Warner, 2008)]. Therefore, there was no significant difference between the mean score (8.74) of the experimental group and the mean score (8.81) of the control group, implying that the experimental and control groups were equally

matched (academically) at the pretest level.

Table 7 – T-test Results for Posttest Scores

Group	Number of Students	Mean	Standard Deviation	Degrees of Freedom	T-Calculated	T-Critical	Decision
Experimental	94	11.70	5.26	184	2.23	1.96	Significant difference
Control	91	9.89	5.77				

Table 7 shows that, at the posttest level, the calculated t value (2.23) was greater than the critical t value (1.96). The effect size, as indexed by $?^2$ [(Warner, 2008)], was .026; this was a medium effect. This implied the existence of a significant difference between the mean score (11.70) of the experimental group and the mean score (9.89) of the control group. The experimental group was statistically superior to the control group in mathematics achievement at the posttest level, suggesting that students taught with simple improvised geometric manipulatives perform better than other students, ceteris paribus.

V. – Conclusion

Based on the finding of this study, persistent poor performance of Nigerian secondary school students in mathematics need not continue indefinitely. There is hope that, with simple, cheap, improvised geometric manipulatives, the situation can be changed for the better. With ordinary cardboard paper, teachers and students can construct and use simple geometric manipulatives to improve teaching and learning of mathematics in Nigeria. By implication, massively harnessing them will enhance the nation's development strides in science and technology.

Acknowledgements

I acknowledge the wonderful research assistance of my son, Toni Aburime, and his wife, Esther Amarachi Nwokeji, toward the successful completion of this paper.

References

Adeniji, J.T. (1998) *WAEC Activities*. A report presented at the STAN Governing Council Meeting, International Conference Centre, Abuja, March 30-31.

Amoo, S.A. (2001) *Curriculum Ideals and Realities for Sustainable Educational Development*. A paper presented at the 14th Annual Conference of the Curriculum Organization of Nigeria, Abuja, Nigeria, September 18-21.

Eze, A.N. (1999) *Practical Approach to Research Methods and Statistics in Education, Management and Social Sciences*. Onitsha-Nigeria: Onwubiko Printing and Packaging Ind. (Nig.) Limited.

Ivowi, U.M.O. (1999) "The Philosophy and Objectives of the Science and Mathematics Curricula at the SS Level". In Akpan, B.B. (Ed.), *Perspectives on Education and Science Teaching: From the Eyes of Uduogie Ivowi*, Ibadan: Science Teachers Association of Nigeria.

Joint Admissions and Matriculation Board (1998) *UME Brochure Guidelines for Admissions for First Degree Courses in Nigerian Universities, 1999-2000 Session*. Lagos: Joint Admissions and Matriculation Board.

National Science Foundation (2002) *Elementary and Secondary Education Curriculum and Instruction.* Arlington, VA: Science and Engineering Indicators.
Warner, R.M. (2008) *Applied Statistics*. California: Sage Publications, Inc.

Mathematical Puzzle

Readers can solve the following puzzles. But, do not send solutions to us.

1. In the grid given below, arrange the numbers 1 to 9 in such a way that each row, column, and 3x3 block contains every number just the once.

2		4		8	5	1		
						9	3	
				4				
		2	7	5				
4			9		6			3
				3	8	2		
				6				
	5	1						
		9	8	7		6		5

2. How can you make two fractions that add up to exactly 1 with all the numbers 0 up to 9 (using each number exactly once)?

3. While I was working with the numbers I found a wonderful three digit number in which the sum of the squares of it's digits was equal to half of that number. Find it.

Solutions to Earlier Puzzles

(in Vol. 1, No. 1, 2008)

2. No, it is impossible. Two diagonally opposite squares on a chess board are of the same color. Thus, when these are removed, the number of squares of one color exceeds by 2 the number of squares of another color. However, each piece of domino covers exactly two squares and these are of different colors. Every placement of domino pieces establishes a one to one correspondence between the set of white squares and the set of black squares. If the two sets have different number of elements, then, by the Pigeonhole Principle, no one-one correspondence between the two sets is possible.

3. The digit 6 can be located in any of the 5 positions; then 8 can be located in 4 positions. Thus the answer is 5 x 4 = 20.

4. The problem is equivalent to finding the number of nonnegative integer solutions of the equation $x_1 + x_2 + x_3 + x_4 + x_5 = 18$.

Interesting Facts about *e*

1. Mathematician Leonard Euler first used *e* in 1736 to mean *that number whose hyperbolic logarithm is equal to 1.*
2. *e* is also known as Euler's number.
3. In 1737, Euler first used the symbol for *pi* to be the ratio of the circumference to the diameter in a circle.
4. In 1777, Euler first used *i* to be equal to the square root of -1.
5. Euler was the first to write five most significant numbers in all of mathematics in one equation
$$e^{(pi*i)} + 1 = 0$$
6. *e* is defined as the limit as *n* approaches infinity of $(1 + 1/n) \wedge n$.
7. *e* = 2.718281828459045 ...
8. *e* can also be defined as the limit as *n* approaches zero of $(1 + n) \wedge (1/n)$.
9. *e* is the base of the exponential function $y = a^x$ such that the slope of the tangent to the curve at the point (0,1) is 1.
10. $$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \ldots$$

Useful Mathematical Symbols

Symbol	What it is	How it is read	How it is used	Sample expression
+	Addition sign Logical OR symbol	... plus ... ... or ...	Sum of a few values Logical disjunction	2 + 6 = 8 ¬(A + B) = ¬A * ¬B
*	Multiplication sign Logical AND symbol	... times ... ... and ...	Product of two values Logical conjunction	2 x 6 = 12 ¬(A * B) = ¬A + ¬B
x	Multiplication sign	... times ...	Product of two values	4 x 2 = 8
·	Multiplication sign	... times ...	Product of two values	4 · 2 = 8
Σ	summation sign	The summation of ...	Sum of many or infinitely many values	$\sum_{n=1}^{\infty} \frac{1}{n}$
$\int$	Integral sign	The integral of ...	integration	$\int x^2\,dx = x^3/3 + c$
$\iint$	double integral sign	The double integral of ...	integration	$\iint f(x,y)\,dx\,dy$
$\iiint$	Triple integral sign	The triple integral of ...	integration	$\iiint f(x,y,z)\,dx\,dy\,dz$
$\oint$	Line integral sign	The line integral of ...	integration	$\oint_{\mathbf{F}} \bullet \mathbf{dx}$
$\iint_S$	Surface integral sign	The surface integral of ...	integration	$\iint_S F(x,y,z)\,dx\,dy$
-	Subtraction sign Minus sign	... minus ... Negative...	Difference of two values, negative number	2 - 6 = -4
±	Plus/minus sign	... plus or minus ...	Expression of range, error, or tolerance	540 kbps ± 10%
•	Dot product sign	... dot ...	Scalar (dot) product of two vectors	**A** • **B** = **B** • **A**
x	Cross product sign	... cross ...	Vector (cross) product of two vectors	**A x B** = - (**B x A**)
Π	Product sign	The product of ...	Product of three up to infinitely many values	$\prod_{n=1}^{\infty} \frac{1}{n}$
^	Carat	... to the power of ...	exponent	2 ^ 4 = 16

!	Exclamation	... factorial	Product of all positive integers up to a certain value	3! = 6
$\sqrt{}$	surd	... root of ...	Algebraic expressions	$z = \sqrt[n]{(x + y)}$
$\sqrt{}$	square root symbol	The square root of ...	Algebraic expressions	$\sqrt{4} = \pm 2$
...	Continuation sign	... and so on up to ... ... and so on indefinitely	Extension of sequence	$\boldsymbol{S}$ = {1, 2, 3, ...}
/	Slash	... divided by ... ... over ...	Division	3/4 = 0.75
÷	Division sign	... divided by ...	Division	3 ÷ 4 = 0.75
%	Percent symbol	... percent ...	Proportion	0.032 = 3.2 %
‰	Per mil symbol	... per mil ...	Proportion	0.032 = 32 ‰
:	Colon, ratio sign	... is to ... ... such that ... ... it is true that ...	Division or ratio, symbol following logical quantifier or used in defining a set	2:4 = 20:40 $\exists x : x > 4$ and $x < 5$ $\forall x : x < 0$ or $x > -1$ $S = \{x : x < 3\}$
|	Vertical line	... such that ... ...it is true that ...	Symbol following logical quantifier or used in defining a set	$\exists x \mid x > 4$ and $x < 5$ $\forall x \mid x < 0$ or $x > -1$ $S = \{x \mid x < 3\}$
::	Double colon	... averaged with ...	arithmetic mean	3 :: 11 = 7
∞	lemniscate	... infinity ... increases without limit	Infinite summations Infinite sequence Limit	$\forall x : x < \infty$
( )	Parentheses	...quantity... ...list... ...set of coordinates... ...open interval	Denotes a quantity, list, set of coordinates, or an open interval	$(x + y) + z$ (a_1, a_2, a_3, a_4) (x,y,z) (3,5)
[ ]	Square brackets	... the quantity ... ... the closed interval ...	Denotes a quantity or a closed interval	$w + [(x + y) + z]$ [2,4]
( ]	Hybrid brackets	... the half-open interval ...	Denotes a half-open interval	(2,4]
[ )	Hybrid brackets	... the half-open interval ...	Denotes a half-open interval	[2,4)
{ }	Curly brackets	... the quantity ... ... the set ...	Denotes a quantity or a set	E = {2, 4, 6, 8, ...}

=	Equal sign	... equals ...	Indicates two values are the same	-(-6) = 6 $2z^2 + 4z - 6 = 0$
$\propto$	Proportionality sign	... is proportional to ...	Indicates two variables change in direct proportion	$x \propto y$
~	Similarity sign	... is similar to ...	Indicates two objects are geometrically similar	$\Delta ABC \sim \Delta DEF$
$\approx$	Approximate equal sign	... is approximately equal to ...	Indicates two values are close to each other	$x + y \approx z$
$\neq$	Inequality sign	... is not equal to ...	Indicates two values are different	$x \neq y$
<	Inequality sign	... is less than ...	Indicates value on left is smaller than value on right	$2 < 4$ $x < y$
$\leq$	Inequality sign	... is less than or equal to ... ... is at most equal to ...	Indicates value on left is smaller than or equal to value on right	$x \leq y$
>	Inequality sign	... is greater than ...	Indicates value on left is larger than value on right	$4 > 1$ $x > y$
$\geq$	Inequality sign	... is greater than or equal to ...	Indicates value on left is larger than or equal to value on right	$x \geq y$
| |	absolute value sign	The absolute value of ...	Distance of value from origin in number line, plane, or space	$\lvert -5 \rvert = 5$
Δ	increment sign, Triangle symbol	the change in ... triangle ...	Indicates a small change, Denotes vertices of triangle	$m = \Delta y / \Delta x$ $\Delta ABC = \Delta DEF$
$\perp$	Perpendicularity symbol	... is perpendicular to ...	Geometry	$L \perp M$
//	Parallel symbol	... is parallel to ...	Geometry	$L \parallel M$
$\angle$	angle symbol	Angle ...	Geometry	$\angle ABC = \angle DEF$
$\exists$	Existential quantifier	For some ... There exists a(n) ...	Logical statements	$\exists x : x > 4$ and $x < 5$
$\forall$	Universal quantifier	For all ... For every ...	Logical statements	$\forall x : x < 0$ or $x > -1$

¬	Logical negation symbol	not ...	Logical statements	¬(¬A) ⟺ A
⟹	logical implication symbol	... implies ... If ... then ...	Logical statements	A ⟹ B
⟺	logical equivalence symbol	... is logically equivalent to ... ... if and only if ..	Logical statements	A ⟺ B
∴	Three dots	... therefore ... ... it follows that ...	Logical statements or mathematical proofs	$x = y$ and $y = z$ $\therefore x = z$
∈	Element-of symbol	... is an element of a set ...	Sets	$a \in A$
∉	Not-element-of symbol	... is not an element of a set ...	Sets	$b \notin A$
⊆	Subset symbol	... is a subset of ...	Sets	A ⊆ B
⊂	Proper subset symbol	... is a proper subset of ...	Sets	A ⊂ B
∪	Union symbol	... union ...	Sets	A ∪ B = B ∪ A
∩	Intersection symbol	... intersect ... ... intersected with ...	Sets	A ∩ B = B ∩ A
∅	Null symbol	The null set The empty set	Sets	∅ = { }
ℵ	Hebrew aleph (uppercase)	Aleph ...	Transfinite cardinal	$\aleph_1 + \aleph_0 = \aleph_1$
°	Degree symbol	... degree(s)	Angular measure Temperature	$\phi = 45°$ $T = +20\ °C$
θ	Greek theta (lowercase)	... theta ...	Angular variable	θ = 90°
ϕ	Greek phi (lowercase)	... phi ...	Angular variable	$\phi = 45°$
λ	Greek lambda (lowercase)	... lambda ...	wavelength Ratio Eigenvalue Lebesgue measure	λ = 70 cm λ = 3:1
μ	Greek mu (lowercase)	micro- (10^{-6})	prefix multiplier	$C = 0.001\ \mu F$
π	Greek pi (lowercase)	... pi ...	General science	π ≈ 3.14159
Ω	Greek omega (uppercase)	... omega ...	Volume of an object Ohms (resistance)	$R_2 = 330\ \Omega$
ω	Greek omega (lowercase)	... omega ...	Transfinite ordinal Angular velocity Period	ω = 36,000 rad/s ω = 1/60 s

$\mathbb{N}$, $\boldsymbol{N}$	Enhanced or bold *N*	The set of natural numbers	Number theory Set theory	$\mathbb{N}$ = {0, 1, 2, 3, ...}
$\mathbb{Z}$, $\boldsymbol{Z}$	Enhanced or bold *Z*	The set of integers	Number theory Set theory	$\mathbb{Z}$ = {0, 1, -1, 2, -2, 3, -3, ...}
$\mathbb{Q}$, $\boldsymbol{Q}$	Enhanced or bold *Q*	The set of rational numbers	Number theory Set theory	$\mathbb{Q}$ = {*a*/*b* | *a* and *b* are in $\mathbb{Z}$}
$\mathbb{R}$, $\boldsymbol{R}$	Enhanced or bold *R*	The set of real numbers	Number theory Set theory	What is the cardinality of $\mathbb{R}$?

ISSN 0974 – 3340

SUTRA

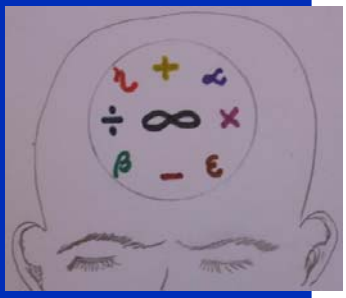
International Journal of

Mathematical Science Education

VOLUME 2, ISSUE 2 JULY-DECEMBER 2009

Published by

Technomathematics Research Foundation, India



EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India.*

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain.*
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK.*
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India.*
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India.*
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India.*
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India.*
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India.*
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia.*
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

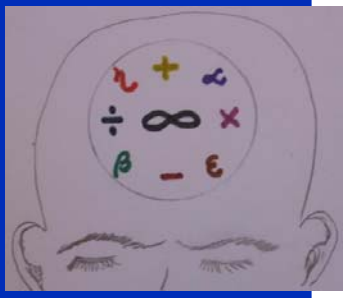
Copyright ©2009 *Technomthematics Research Foundation.*

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.

Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.

Sutra: International Journal of Mathematical Science Education (ISSN **0974 – 3340**) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education



CONTENTS (Vol. 2, No. 2, 2009)

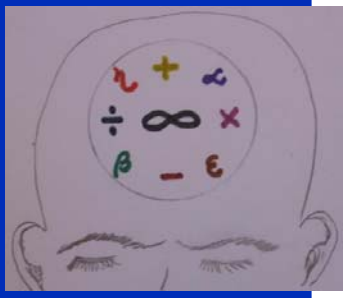
SOME APPLICATIONS OF EULERIAN GRAPHS. 1- 10
Abdul Samad Ismail , Roslan Hasni and K. G. Subramanian

ENGAGING STEM STUDENTS IN RESEARCH AND PRESENTATION ACTIVITIES. 11 - 23
Rohitha Goonatilake and Rafic A. Bachnak

A NEW METHOD FOR SOLVING QUARTICS. 24- 26
Raghavendra G. Kulkarni

A NUMBER THEORETIC IDENTITY. 22 - 28
Amiya Bhushan Sarkar

Sutra: International Journal of Mathematical Science Education



EDITORIAL (Vol. 2, No. 2, 2009)

I am glad to present a new issue of the Sutra journal. This issue contains four papers. Interesting applications of graph theory are discussed in the first expository paper. Second paper presents results from a project conducted at Texas A&M International University. Next two short papers present a method for solving general quartic equation, and a number theoretic identity.

I hope that readers will enjoy reading this issue.

R. Akerkar

Editorial Coordinator

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 2, 1 – 10, 2009

SOME APPLICATIONS OF EULERIAN GRAPHS

Abdul Samad Ismail , Roslan Hasni and K. G. Subramanian

School of Mathematical Sciences,
Universiti Sains Malaysia,11800 Penang,Malaysia
samafa70@yahoo.com, hroslan@cs.usm.my, kgs@usm.my

The genius Swiss Mathematician Leonhard Euler who was a prolific contributor to several areas of Mathematics is considered as the inventor of the concept of a graph. Graphs have proved to be very useful in modeling a variety of real-life situations in many disciplines. This article is intended for the attention of young readers, uninitiated in graph theory and gives an introductory discussion of certain well-known application problems that involve graphs in their solution and in particular the role of Eulerian graphs starting from the problem of Konigsberg seven bridges to the current problem of DNA fragment assembly. We also point out connections of Eulerian circuits with drawing certain floor designs, known as "kolam".

Keywords: Graph; Eulerian circuit; Eulerian graph

1. Introduction

Topics under the broad title of "Discrete Mathematics" are intended to provide the mathematical foundation for pursuing courses relating to computer science besides giving a mathematical background for solving problems in many application areas. Graph theory is one such topic which has found use in a variety of application problems. The birth of Graph theory can be attributed to the work of the famous Swiss mathematician Leonhard Euler (1707 – 1783) in solving the problem of "Seven bridges of Konigsberg". We discuss this problem in a later section. The method used by Euler in solving this problem gave rise to the notion of 'graph', which essentially is a discrete structure useful for modelling relations among objects. The development of the subject of graph theory has therefore been phenomenal with the subject drawing from and contributing to many other disciplines of study. In this article we discuss the notions of a graph, Eulerian graph and certain application problems that involve Eulerian graphs, starting from the problem of Konigsberg seven bridges to the current problem of DNA fragment assembly, to draw the attention of the young researchers uninitiated in this area of study. We also bring out some connections of Eulerian circuits with certain floor designs, known as "kolam".

2. Acquaintance Problem and Graphs

A group leader of a company was addressing in a training programme, a group of 21 newly recruited trainees of the company, who have assembled for the programme. The

leader who was fond of numbers was making a remark as follows: Among the 21 trainees assembled here it could be that there are trainees who know each other, trainees who know all others, trainees who do not know each other as well as trainees who do not know anyone in the group. But there will be a trainee in this group who already knows an even number (including the even number zero) of other trainees. The claim of the leader looks a little puzzling but it is true. What do we do to prove this claim? One way is to create a list, for each of the 21 persons, of other trainees in the group already known to them. The information obtained can be processed, although it could be a little cumbersome, in order to find out for each of the 21 trainees, the number of other trainees already known to them. This will reveal the truth of the claim.

On the other hand, it is interesting to note that this problem could be discussed with the use of graphs. We shall illustrate this with a smaller number of 7 trainees in the group, say Hussain, Johnson, Kamalesh, Khairul, Prasanna, Qadir, Ramesh. For convenience, we code their names as H,J,K1,K2,P,Q,R. Suppose H knows R (even before meeting R in this training programme); J knows K1,P,Q; K1 knows J,K2,Q; K2 knows K1; P knows J; Q knows J,K1; R knows H. It is clear that H knows only one other trainee; J knows three; K1 knows three; K2 knows 1; P knows 1; Q knows 2; R knows 1. We shall call these numbers as 'acquaintance numbers'. Thus the acquaintance numbers of H,J,K1,K2,P,Q,R are respectively 1,3,3,1,1,2,1. The claim of the leader is verified, since (at least one of them) Q knows an even number of other trainees.

The situation can be modeled by a 'graph'. For each trainee we associate a point (indicated by a small circle) in the plane. We label the points as H,J,K1,K2,P,Q,R. We join two points by a line (straight or curved) if the two trainees represented by the points 'know' each other. The resulting diagram which is an example graph, is in Figure 1. The number of lines at a point gives the acquaintance number of the corresponding person.

Consider another situation as follows: Suppose H knows only R; J knows only K1; K1 knows only J; K2 knows none; P knows only Q; Q knows only P; R knows only H. Then the corresponding graph is shown in Figure 2. The acquaintance numbers are 1,1,1,0,1,1,1. Still one of these numbers namely 0 is even, again verifying the claim. Note that if the number of trainees in our discussion is 6 or 8 instead of 7 , the claim need no longer be true. For example, if there are only 6 trainees H,J,K1,K2,P,Q such that H knows only J; J knows only H; K1 knows only K2; K2 knows only K1; P knows only Q; Q knows only P, then the acquaintance numbers are 1,1,1,1,1,1, all of which are odd.

So, for the claim to be true the number of trainees in the group should be odd. This fact can be proved as a consequence of a basic theorem, known as "Handshaking Theorem" in graph theory, which we shall discuss a little later.

Thus a graph is a discrete structure that gives a representation of a finite set of objects and certain relation among some (or all) objects in the set. We shall now express the notion of a graph and certain terms related to graphs in a little more rigorous way.

A graph G = (V,E) consists of a finite set V of elements, called vertices (also called points or nodes) and another finite set E of pairs of elements of V of the form (u,v) (also written as uv or vu). The element (u,v) is called an edge (also called line). In a diagrammatic representation, we denote a vertex by a small circle (or by a small dot) and an edge (u,v) by a line joining the points representing the vertices u and v. For example, in Figure 3, a graph is shown with 5 vertices u,v,w,x,y and 5 edges (u,x), (u,w), (x,w), (x,y), (v,y). Note that in a graph there could be more than one edge between two vertices which are referred to as multiple edges between the vertices or there could be an edge (called loop) joining a vertex with itself. Such a graph is called a multigraph. In Figure 4, a graph with multiple edges and loops is shown.

The number of edges at a vertex is called the degree of the vertex. The degree of the vertex u in the graph in Figure 3 is therefore 2. The degrees of v, w, x, y are respectively 1, 2, 3, 2. If we add all these degrees we have the sum = 2+1+2+3+2 = 10. This number is twice the number of edges, namely 5. In fact this is true in any graph and the result is referred to as "Handshaking Theorem". Note that in the acquaintance graph in Figure 1, the number of edges is 6 and the sum of the degrees (or the acquaintance numbers) equals 1+3+3+1+1+2+1 = 12 and in Figure 2, the number of edges is 3 and the sum of the degrees equals 1+1+1+0+1+1+1 = 6.

Handshaking Theorem: In any graph, the sum of the degrees of the vertices is twice the number of edges.

If we denote the degree of a vertex v in a graph G = (V,E) by d(v) and if G has q edges, then this theorem can be expressed as follows:

$$\Sigma_{v \varepsilon V} d(v) = 2q.$$

Note that this theorem is a consequence of the fact that an edge (which is not a loop) is incident with two vertices, thus contributing degree one to each vertex. So while counting the degrees of vertices, such an edge will get counted twice, once at each end of the edge. (A loop is incident with a vertex two times and so contributes degree two to this vertex). In a graph a vertex with even degree is called an even vertex and a vertex with odd degree as an odd vertex. The sum of the degrees of even vertices is an even number since we always have an even integer when we add any number of even integers. But only the sum of an even number of odd integers is an even integer. As a consequence of Handshaking Theorem, the number of odd vertices in a graph has to be even. In fact in a graph with an odd number of vertices, the vertices of odd degree will be even in number and so at least one vertex must have even degree. This is the reason why the claim of the

group leader in the acquaintance problem with an odd number of trainees that there is at least one trainee who knows an even number of other trainees, is true. For example, in the graph with seven vertices in Figure 1, there are an even number (six) of odd vertices that correspond to trainees who know an odd number of other trainees.

3. Diagram Tracing and Eulerian graph:

A popular old game that entertains children runs as follows: Can you trace with a pencil a diagram of points (with the small circles representing the points) and lines as shown in Figures 5a and 5b? The condition is that the diagram is to be traced beginning at a point and on completion end at the same point but the pencil should not be lifted till the diagram is completely traced and a line in the diagram should not be retraced (i.e. can be traced only once). A curious child will certainly try to find the answer by trial and error and will arrive at the conclusion after sometime that it is not possible in diagram in Figure 5b but it is possible in the diagram in Figure 5a. The question of whether it is possible to trace such a diagram, can be quickly answered if the concept of Eulerian graph is known. Indeed the diagram in Figure 5a is Eulerian whereas the diagram in Figure 5b is not.

We now introduce the concepts of path and circuit in a graph to enable us to describe the notion of an Eulerian graph in a little more rigorous way.

If e = xy is an edge in a graph, then x is called the start vertex and y, the end vertex of e. A path P (or u-v path P) in a graph is a sequence of edges so that the end vertex of an edge in the sequence is the start vertex of the next edge in the sequence and the path begins in the vertex u and ends in the vertex v. If vertices u and v are the same, then the path is called a circuit (some call it cycle). For example in the graph in Figure 6, (a,b)(b,c)(c,b)(b,c)(c,d) is a path and (a,b)(b,c)(c,b)(b,c)(c,a) is a circuit. Vertices and/or edges can be repeated in a path or in a circuit. (A path is called a walk by some authors. Due to the diversity of people who use graphs for their own purpose, the naming of certain concepts has not been uniform in graph theory).

An Eulerian path in a graph is a path which uses all the edges of the graph but uses each edge exactly once. An Eulerian circuit is a circuit which has a similar property. Note that in an Eulerian path or an Eulerian circuit, a vertex can be visited more than once but not an edge ! An Eulerian circuit begins and ends at the same vertex. A graph that has an Euler circuit is known as an Eulerian graph. For example in the graph in Figure 6, (a,b)(b,c)(c,e)(e,d)(d,c)(c,a) is an Eulerian circuit and hence the graph in Figure 6 is an Eulerian graph.

4. The Problem of Seven Bridges

The year 1736 when Euler solved the problem of seven bridges of Königsberg is taken to mark the birth of graph theory. The seven bridges problem is a well known problem that can be stated as follows: The Pregel river in the town of Königsberg divided it four land regions with a central island. These four land regions were connected by seven bridges as shown in the diagram in Figure 7a where the land areas are denoted by the letters A,B,C,D. It is said that the people of the town used to ponder over the question of whether it is possible to start in a land area and make a walk through the bridges visiting each bridge exactly once and returning to the starting point. In solving this problem of "Seven bridges", Euler developed an abstract approach which is considered to be related to the "geometry of position". Represented in terms of a graph E_g (Figure 7b), the four land areas A, B, C, D are the vertices and the seven bridges are the edges of the graph. The vertices A and B are joined by multiple edges (two edges), so that the graph E_g is a multigraph! The "Seven bridges problem" can be expressed in graph-theoretic terms as follows: Is there an Eulerian circuit in the graph E_g? The answer from Euler's 1736 paper to this question is NO! . This is stated as an important theorem in the study of Eulerian graphs.

Theorem on Eulerian graphs: A connected graph with two or more vertices is an Eulerian graph (ie. has an Eulerian circuit) if and only if each vertex of the graph has even degree.

Note that the necessary part of the theorem is based on the fact that, in an Eulerian graph, everytime a circuit enters a vertex through an edge it exits the vertex through another edge, thus accounting for an even degree two at the vertex. At the starting vertex, the circuit initially exits and finally enters when the circuit is completed.

In the graph E_g in the Seven bridges problem, all the four vertices have odd degree. So it cannot have an Eulerian circuit.

5. Chinese Postman Problem

All vertices of a graph need not be of even degree and so a graph may not be Eulerian. But if a graph has all except two vertices of even degree then it has an Eulerian path which starts at one of the odd vertices and ends at the other odd vertex. A graph having an Eulerian path but not an Eulerian circuit is called semi-Eulerian. For example in the graph in Figure 8, (a,b)(b,c)(c,d)(d,b)(b,e)(e,d)(d,f) is an Eulerian path and hence the graph in Figure 8 is semi-Eulerian.

In a graph G_p that models streets and street corners in a town, with the street corners as the vertices and the streets as the edges, some one starting in a street corner and walking along the streets one after another can end up in a street corner, giving rise to a path in the

graph. Note that walking along a street more than once from one corner to another, corresponds to repeating an edge or a vertex in the graph.

Suppose that a postman has to deliver letters to the residents in all the streets of a village. Assume that the village is small enough for the postman to be assigned this task everyday. If the graph G_p that represents the streets and street corners as mentioned above, is semi-Eulerian so that there is an Euler path in the graph G_p, then this path gives rise to a route following which the postman can start in a street corner, deliver letters going through every street exactly once. This is a desirable situation for the postman. But if no such Euler path exists in G_p, then the postman may have to repeat some of the streets. This problem is called the Chinese postman problem in honour of a Chinese mathematician Meigu Guan who proposed this problem.

6. DNA fragment assembly

DNA (deoxyribonucleic acid) is found in every living organism and is a storage medium for genetic information. A DNA strand is composed of bases which are denoted by A (adenine), C (cytosine), G (Guanine) and T (thymine). The familiar DNA double helix arises by the bondage of two separate strands with the Watson-Crick complementarity (A and T are complementary; C and G are complementary) leading to the formation of such double strands.

DNA sequencing and fragment assembly is the problem of reconstructing full strands of DNA based on the pieces of data recorded. It is of interest to note that ideas from graph theory, especially Eulerian circuits have been used in a recently proposed approach to the problem of DNA fragment assembly. We do not enter into the details but only mention that this brings out the application of graph theory in the field of bioinformatics. More elaborate details on this topic can be found in Kaptcianos (2008) and Pevzner et al (2001).

7. Floor Designs and Eulerian graphs

Traditional interesting floor designs, known as “kolam”, are drawn as decorations in the floor and in large sizes and in interesting shapes, during festivals and weddings, with the drawing done with rice flour or rice paste especially in South India. See for example Nagata and Robinson (2006), Siromoney (1978) and Siromoney et al (1974). Generally, in drawing a kolam, first a suitable arrangement of dots is made and then lines going around the dots are drawn. An example kolam is shown in Figure 9 where the dots are ignored. One of the types of kolam, is known as ‘kambi kolam’ (the Tamil word kambi meaning wire). In this type, the kolam is made of one or more of such ‘kambi’s . The kolam in Figure 9 is of this type and is made of a single ‘kambi’ whereas the kolam in Figure 11 is made of three ‘kambi’s.

A kolam drawing can be treated as a special kind of a graph with the crossings considered as vertices and the parts of the kambi between vertices treated as edges. The only restriction is that unlike in a graph, these edges can not be freely drawn as there is a specific way of drawing the kolam. The single kambi kolam will then be an Eulerian graph with the drawing starting and ending in the same vertex and passing through every edge of the graph only once. In Figure 10, the kambi kolam of Figure 9 is shown as a graph with vertices (indicated by small thick dots) and edges and this graph is Eulerian as every vertex is of degree 4. Note that the graph of kambi kolam (with more than one kambi) in Figure 11 will also be Eulerian but the drawing of the kolam in the way it will be done by the kolam practitioners will not be giving rise to the Eulerian circuit. On the other hand in the case of single kambi kolam it is of interest to note that the drawing of the kolam will give a tracing of an Eulerian circuit in the corresponding graph.

8. Conclusion

With an intention to bring to the notice of young readers, some of the applications of graphs, especially Eulerian graphs are pointed out in this article. There are a number of other interesting applications where graphs have found their use. An account of some of the other applications is given in Ismail (2009). The authors acknowledge consulting the books (Akerkar and Akerkar (2004), Arumugam and Ramachandran (2001), Fleischner (1990), Rosen (2007)) in the preparation of this article. The interested reader is encouraged to refer to them and the references in them for more details.

References

Akerkar, R. A. and Akerkar, R. R., (2004) *Discrete Mathematics*, Pearson Education (India).

Arumugam, S. and Ramachandran, S., (2001) *Invitation to graph theory*, Scitech Publications (India) Pvt. Ltd., Chennai (India).

Fleischner, H. (1990) *Eulerian Graphs and Related Topics*, Part 1, Vol. 1. Annals of Discrete Mathematics 45, Elsevier Science Publishers B. V., Amsterdam-New York.

Ismail A.S. (2009) *An introduction to Eulerian and Hamiltonian graphs*, M.Sc Project Report, PPSM, USM.

Kaptcianos, J. (2008) A graph-theoretical approach to DNA fragment assembly, *American Journal of undergraduate Research*, 7(1).

Nagata, S. and Robinson, T., (2006) Digitalization of kolam patterns and tactile kolam tools, In '*Formal Models, Languages and Applications*' (K.G. Subramanian, K. Rangarajan and M. Mukund, eds.), Series in Machine Perception and Artificial Intelligence, Vol. 66, 354-363.

Pevzner, P.A., Tang, H. and Waterman M.S. (2001) An eulerian path approach to DNA fragment assembly, *Proc. Natl. Acad. Sci. USA.* 98, 9748-9753.

Rosen, K.H., (2007) *Discrete Mathematics and its Applications*, McGraw-Hill International Edition.

Siromoney, G. (1978) South Indian kolam patterns, *Kalakshetra Quarterly*, 1, 9-14.

Siromoney, G., Siromoney, R. and Krithivasan, K. (1974) Array grammars and kolam, *Computer Graphics and Image Process.*, 3, 63-82.

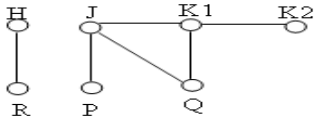


Fig. 1 Graph of an acquaintance of trainees

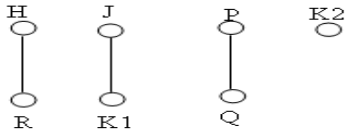


Fig. 2 Graph of another acquaintance of trainees

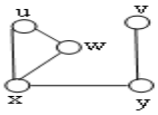


Fig. 3: A graph on five vertices

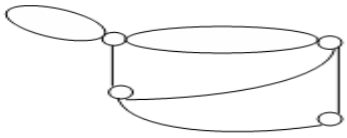


Fig.4 : A multigraph

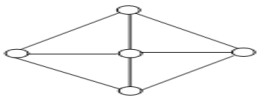


Fig. 5b: Diagram that cannot be traced

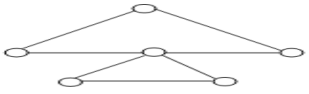


Fig. 5a : Diagram that can be traced

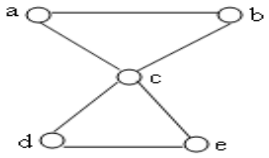


Fig. 6: An Eulerian graph

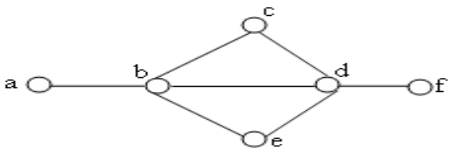


Fig. 8: A semi-Eulerian graph

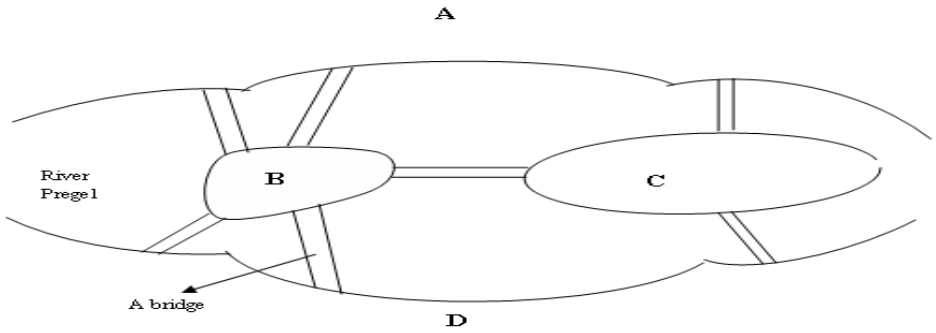


Fig. 7a: Seven bridges of Konigsberg - a sketch

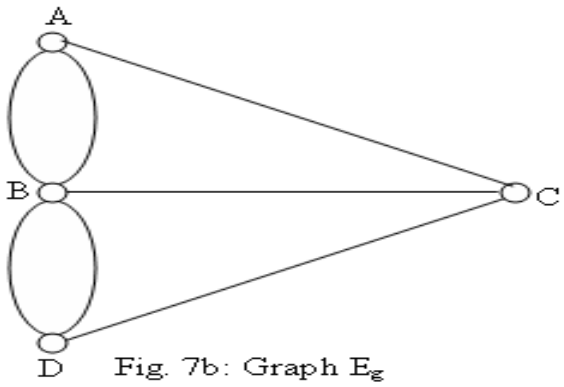


Fig. 7b: Graph E_g

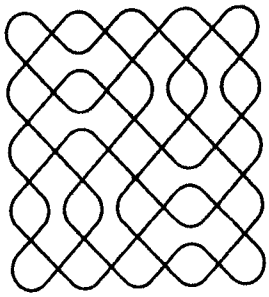


Fig. 9: A single kambi kolam

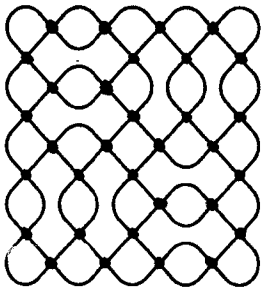


Fig.10: Graph of Kolam in Fig.9

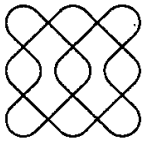


Fig. 11: A kolam made of three Kambis

Brief Biographical sketch of the authors:

Abdul Samad Ismail is a Mathematics teacher in Penang Free School, Penang, Malaysia. He did his part-time post-graduate studies in the School of Mathematical Sciences, Universiti Sains Malaysia during 2006-2009 and he did a project on Eulerian graphs under the supervision of Prof. K.G. Subramanian.

Roslan Hasni is a senior lecturer in the School of Mathematical Sciences, Universiti Sains Malaysia. He was awarded PhD degree in the year 2005 from Universiti Putra Malaysia. His area of research is graph theory, especially graph colorings and domination in graphs.

K.G. Subramanian is a visiting Professor from 2007 at the School of Mathematical Sciences, Universiti Sains Malaysia, after retiring in 2006 from the Department of Mathematics, Madras Christian College, Chennai, India. His main area of research is two-dimensional picture array languages. His PhD supervisor is the eminent Indian formal language-theorist Prof. Rani Siromoney.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 2, 11 – 23, 2009

ENGAGING STEM STUDENTS IN RESEARCH AND PRESENTATION ACTIVITIES

Rohitha Goonatilake and Rafic A. Bachnak
Department of Engineering, Mathematics, and Physics
Texas A&M International University
5201 University Boulevard
Laredo, Texas 78041-1900, USA
E-mails: harag@tamiu.edu and rbachnak@tamiu.edu

Abstract

College faculty often attempt to find innovative methods to facilitate student learning and retention. One way of accommodating this is to involve students in research projects that offer students the opportunity to present at local conferences or educational forums, as attending regional conferences is rather expensive. This can be done in many ways. Student presentations, for example, can be part of the evaluation of course grades. But they can also be a part of a large educational project that contributes to student learning, success, and retention. This paper outlines a project undertaking conducted at Texas A&M International University (TAMIU) with an emphasis on the extent of such educational endeavors. Student self-evaluation results are used as an indication of how successful the project was in achieving desired outcomes.

Keywords: Students, presentations, STEM disciplines, self-evaluation, Hispanic

1. Introduction

Oral presentations provide a learning opportunity for students to develop their communication skills. Most importantly, they serve to strengthen students' knowledge about a topic of interest or subject that they wish to pursue. A key feature of a high quality learning experience for students in an academic institution is establishing a scholarly culture that includes integration of research into the learning process in which students are encouraged to construct their own meanings and deep understandings of

the subject materials. This process makes students engage in critically constructing knowledge and does not merely depend on instructions being provided to them with access to the contents. The learning experience therefore is facilitated by sustained interaction between the faculty and students in the classroom and beyond.

On the other hand, learning keeps students and their teachers in touch with understanding and achievement of subjects on a continuous basis, thus allowing students to know what specific actions they must take to improve the learning everyday of their lives. Depending upon the nature of the task, these skills may be sometime very generic in nature, such as public speaking or more specific to STEM (Science, Technology, Engineering, and Mathematics) students involved in research projects. For example, a presenter is able to communicate the findings of primary research directly to the audience using graphs and charts to present information and discuss at length the research project, including responses to comments and questions from the audience. While this interaction helps the presenter in developing listening and reasoning skills, it also helps students in the audience learn from observing and asking questions, one aspect of presentations that is generally neglected. Questions and responses are normally instantaneous, spontaneous, and well thought out. This undertaking is mainly carried out under the auspicious of STEM Research Experiences to Enhance Retention (STEM-REER), which is part of STEM Recruitment, Retention, and Graduation (STEM-RRG). STEM-RRG is an innovative recruitment and retention program, funded by the US Department of Education, to increase the number of Hispanic and other minority students pursuing degrees in STEM fields at TAMIU. Studies have suggested that involving undergraduate students in research has several benefits, including attaining a higher level of competence, understanding the methods and process of research, making informed judgments about technical matters, working in teams to solve complex problems, improving their oral and written communication skills, and pursuing graduate degrees.

2. Objectives of STEM-REER

This project offers real-world learning experiences to outstanding students in STEM majors at TAMIU. The project is aimed primarily at students who have the desire and talent but are constrained by various environmental factors. The program awards 10 internships and 20 research

assistantships to highly qualified students. Interns are placed in private industry or government positions and mentored by the client organization. Research assistants are placed with STEM faculty members doing research in computer science, mathematics, biology, physics, or engineering. The goals of this project are: 1) To offer real-world learning experiences to help students complete undergraduate degrees in engineering, mathematics, biology, chemistry, and physical science; 2) To increase the retention of talented minority students in the STEM disciplines and accelerate their pace toward obtaining their undergraduate degrees; 3) To improve students' employability through career planning and internships in industry; and 4) To improve graduate school opportunities through participation in undergraduate research with a faculty mentor. Interns and research assistants are required to: a) Meet regularly with a faculty member, who will serve as the students' mentor for the entire period of the program; b) Participate in a monthly colloquium and share experiences with students and faculty; c) Attend on-campus seminars offered by faculty, representatives and experts from industry; d) Prepare and submit a progress report at the end of each semester; and e) Become active members of the Society of Engineering at TAMIU (SET), the Mathematics Society of TAMIU (MST), or the TAMIU Biology Club.

The project also provides students with a basic groundwork on management, operations, and research techniques, focusing on problem formulation and model construction. Finding a quantitative solution, interpretation and performing sensitivity analysis for deterministic problems using spreadsheet capabilities is generally common. At the end of their presentations, they must be able to show evidence of their ability to plan, execute, manage, control, and report on the project independently or in group setting. Finally, if someone is unable to carry out the tasks expected of the project, the presenter needs to seek mentoring and further faculty intervention.

The reasons for a presenter to miss making a planned presentation can vary from logistic nature to a simple avoidable circumstance. This project has two similar cancellations altogether. In one, the presenter could not make it on time and other was the presenter did not receive the e-mail communication informing the time and day of the presentation. The latter

has been subsequently made a classroom presentation. There was a case where the presenter had to change the topic of the presentation.

3. Methodology and Analysis

For the analysis, a group of 20 presentations (24 presenters) at The 2009 Lamar Bruni Vergara University Conference at TAMIU, held from April 20-23, 2009, was examined [LBV University Conference]. Of 17 oral presentations and 3 poster presentations, 10 were graduate student presentations, 9 were STEM-RRG presentations, 3 were honors student presentations, and 2 were undergraduate class project presentations done by course re-takers. All STEM fields were represented ranging from mathematics (50.00%), biology (20.83%), statistics (8.35%), engineering (12.50%), computer science (4.16%), to management information systems (4.16%) as appears in Tables 1 and 2, for those who participated in the self-evaluation (see the appendix). In addition, there were seven other STEM-RRG presenters in Table 3 who did not participate in the self-evaluation. All STEM-RRG presentations were from the projects the student research assistants undertook under this program and supervised by STEM faculty [STEM RRG Project]. This faculty member essentially mentors these students how to conduct research, documenting and presenting them. Most of the cases, research topics are selected by the students in consultation with the faculty member. The honors students on the other hand need to undertake projects related to the course and then present at a local venue before the final project report is submitted for honors credits. These project assignments will also further facilitate students' improved understanding of conceptual knowledge as they were assigned written research papers, oral presentations or poster presentations. The process of explaining to themselves and to others during the practice sessions and in the actual presentation develops and constructs conceptual and extended knowledge of the topic [Kitto (2008)].

4. Tables and Figures

Tables 1 and 2 provide the breakdown of student presentations by major, classification, and title of presentation or poster for those who participated in the self-evaluation.

Table 1: Programs, classifications, and titles for each STEM-RRG presentations

Student Major, Classification	**Title of Presentation/Poster**

MIS, Graduate	Software as a Service
MIS, Graduate	HY WIRE: A Revolution in Mobility
MIS, Graduate	Logic Gates (Poster)
MIS, Graduate	Trigonometry: Pythagorean Theorem (Poster)
Engineering, Junior	Rates of Change: Position of a Falling Object (Poster)
Engineering, Junior	Topography of an Object: Detection and Display of Software and Hardware
Biology, Freshman	The Impact of Natural Selection on the Evolutionary Process of Species
Biology, Freshman	The Study of Heat Shock Proteins and Adaptive Thermo Tolerance
Biology, Junior	Smallpox

Table 2: Programs, classifications, and titles for each course-related presentation

Student Major, Classification	**Title of Presentation/Poster**
Mathematics, Graduate	Nonlinear Systems of Differential Equations: Dynamical Systems
Mathematics, Graduate	Equilibria in Nonlinear Systems of Differential Equations
Mathematics, Graduate	Global Nonlinear Techniques: Nullclines
Mathematics, Graduate	Nonlinear Systems of Differential Equations: Closed Orbits and Limit Sets
Mathematics, Graduate	Nonlinear Linear Systems of Differential Equations: Applications in Circuit Theory
Mathematics, Graduate	Nonlinear Linear Systems of Differential Equations: Applications in Mechanics
Mathematics, Senior	Divorce and Marriage Rates: Local vs. National Data

Biology, Sophomore	A Statistical Overview of the Service-related Activities by the Doctors Hospital of Laredo
Psychology/Statistics, Freshman	A Statistical Overview of Over the Counter Drug Abuse
Mathematics, Senior	Sudoku
Mathematics, Senior	Blackjack (a class presentation)

Table 3 is the list of programs, classifications, and titles for the other seven presentations that consist of topics from computer science and biology for those presenters who did not participate in the self-evaluation process.

Table 3: Programs, classifications, and titles for each other STEM-RRG presentation

Student Major, Classification	**Title of Presentation**
Computer Science, Senior	Peer to Peer Networks
Biology, Senior	Assessment of Water Distribution Changes in Human Cortical Bone by NMR
Biology, Junior	The Characteristics of Simulated and Functional Disuse on Mouse Bone by Nuclear Magnetic Resonance
Biology, Junior	The Geodynamo Theory and How It Affects the Earth
Biology, Senior	Survey of the Presence of Lectin Activity in South Texas Plants
Biology, Senior	The Use of the Tandem Inverted Repeat System to Screen Carbonic Anhydrases: RNAi Transformants in Chlamydomonas Reinhardtii
Biology, Senior	Patterns and Distribution of Mealy Galls in South Texas

All presenters in Tables 1 and 2 were requested to participate in the self-evaluation. The results derived from 24 self-evaluation instruments are summarized in Figures 1-6.

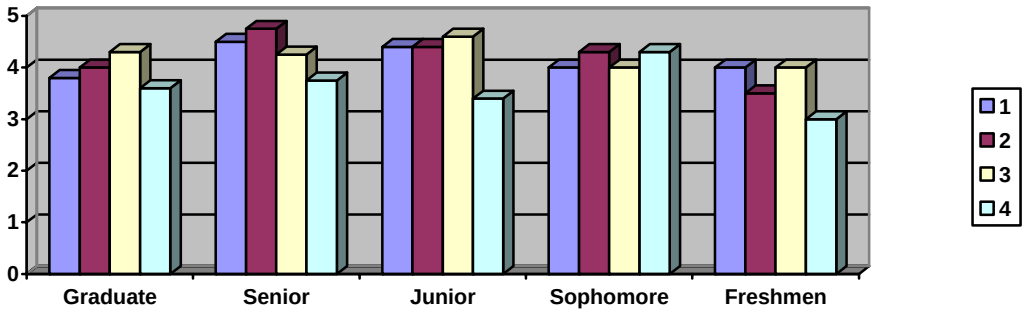


Figure 1: The student self evaluation of presentations: 1-overall presentation, 2-extent of learning outside the classroom, 3-extent of thoroughness, and 4-difficulty of presentation on a scale of 1-5 for graduate, senior, junior, sophomore, and freshman students

Freshmen encountered least amount of difficulties in their presentation; seniors exhibited extensive learning activities outside the classroom, and juniors expressed thoroughness and satisfaction with their overall presentations as were evident from Figure 1.

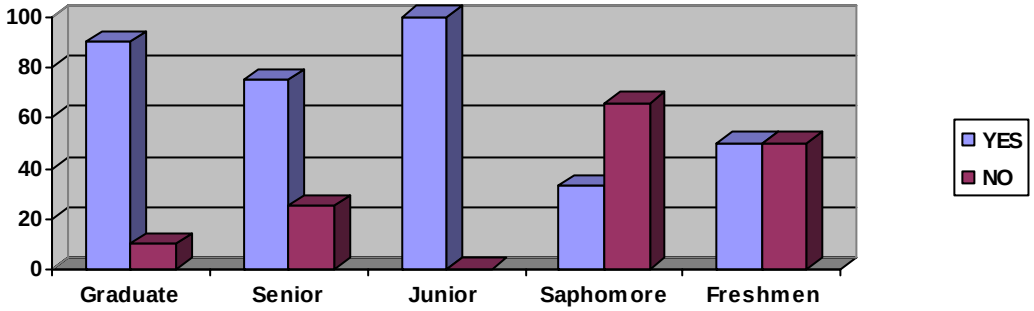


Figure 2: The extent of preparation through rehearsals before actually presentation

Most graduate and junior students indicated that they have rehearsed the presentation as per Figure 2. About fifty percent of freshmen had rehearsed their presentations.

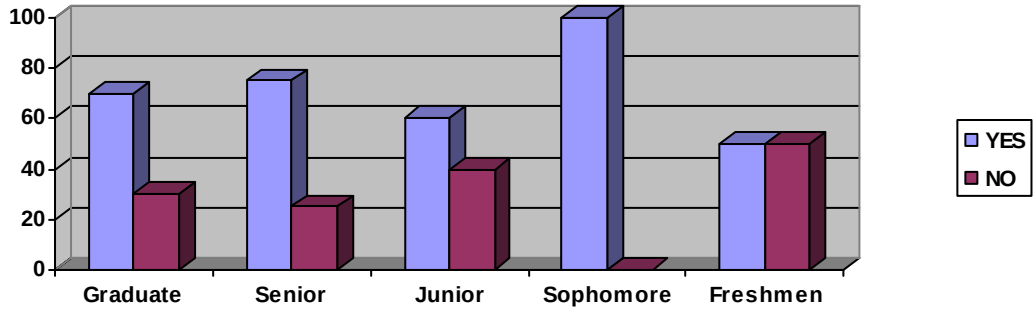


Figure 3: The extent of participants seeking outside help

According to Figure 3, sophomore students sought the most outside help in getting ready for the presentations. About fifty percent of freshmen sought outside help for their presentations.

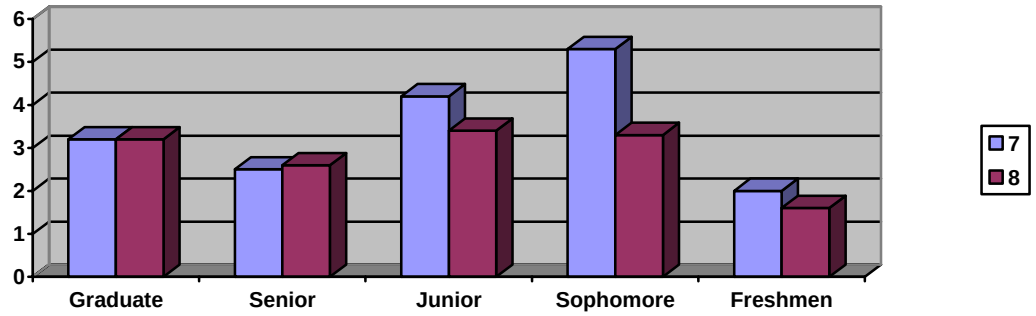


Figure 4: The extent of time spent preparing (7) and gathering information (8)

The graduate students have spent equal amount of time preparing and gathering information for their respective projects as depicted in Figure 4.

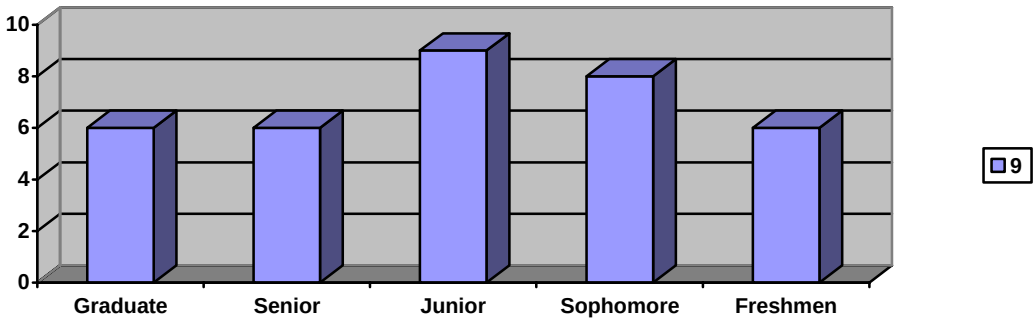


Figure 5: The average number of sources consulted for the delivery of the presentation

Among all presenters, the juniors and sophomores have consulted more sources in preparation for their presentations according to Figure 5.

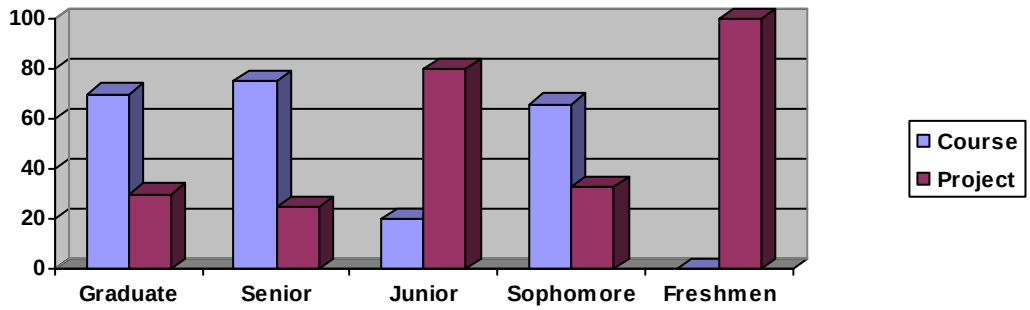


Figure 6: The percentage of presentations that are course-related or project-related topics

Figure 6 shows that most of the presentations by juniors and freshmen were project-related while the presentations by seniors and graduates were course-related.

5. Discussion

Adequate cognitive flexibility in these forums by taking on multiple cognitive perspectives helps the learner to absorb new information and particularly the faculty should be flexible as such both of them are trying to understand the new topic for furtherance of the research topic in hand or assigned. More importantly, this means to transparently communicate to

students some of one's personal understanding in response to their actions, whenever this is considered supportive in the argument. For example, the faculty mentor might genuinely comment on a student's presentation before it is actually presented. There were occasions in the past where some students felt that they were bored by another presentation as it was almost a repetition of the materials they have previously learned or have seen [Boud (1995)].

Analysis and utilization of literature provided the presenter an opportunity to learn beyond the scope of the course materials for his or her advancement. In preparation for this project, a presenter may have reviewed as much as possible, published articles and reports that address the outlines of the project. They have definitely cited some information from them. The faculty involvement is highly critical as the benefits of undergraduate research defined as the engagement of undergraduate students in authentic research conducted in long and intensive summer programs under the direct supervision of faculty researchers. Project organizers were concerned to establish what is known with any certainty the qualities and effects of research, thus defined. The previously published work has been developed on the basis of research and evaluation, and the completeness and quality of the evidence offered to them in support of hypothesized arguments to participating students in a variety of different formats [Seymour, Hunter, Laursen, and Deantoni (2003)].

A mechanism for rapid assessment must be established to allow the presenter to benefit from the results of the self-evaluation. Self and peer-assessment will have dual roles. One is that the presenter will be able to evaluate in writing their presentation themselves closely. The other is to sort out their strengths and weaknesses for future presentations. The use of posters to accomplish this task is impressive. It not only conveys the main theme of the project undertaking, but also clear and reflective of the findings for audience to easily grasp. The assessment of these projects must meet these aspects. The criteria of a similar assessment can be less dependable for complex and open-end responses. It can be either self or peer-assessment, it must also agree within a specified range, whether expressed as a qualitative grade or as ranking, award points or the higher grade if it is done for a class project [Motschnig-Pitrik and Holzinger (2002)].

6. Conclusions and Future Work

Participants in this project believe that presenting in the conference was a quite learning experience for them. Several participants were the recipients of awards. One oral presentation received master's level second place and two presentations received first place and second place undergraduate awards. As we concluded initial results unique to this study, these results are subject to further investigations before validation.

Acknowledgements

STEM-RRG is funded by the U.S. Department of Education (Award # P031C080083). More information about this grant is found at the website: http://www.tamiu.edu/~rbachnak/STEMRRG/index3.html. The work done by Sravan Kumar Madadi and Sandeep Theeda, two students assigned to this project is acknowledged. The authors want to thank faculty mentors, all presenters, particularly those who participated in the self-evaluation process.

References

Boud, D. (1995): Enhancing Learning through Self Assessment, (London: Kogan Page)

Kitto, K. L. (2008): Developing and Assessing Conceptual Understanding in Materials Engineering Using Written Research Papers and Oral Poster Presentations, 38th ASEE/IEEE Frontiers in Education Conference, October 22-25, Saratoga Springs, NY, pp. 1-6

LBV University Conference: http://www.tamiu.edu/gradschool/lbvconference/

Motschnig-Pitrik, R. and Holzinger, A. (2002): Student-Centered Teaching Meets New Media: Concept and Case Study, Educational Technology and Society 5 (4)

Seymour, E.; Hunter, A-B; Laursen, S. L.; and Deantoni, T.(2003): Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study Ethnography & Evaluation Research, Center to Advance Research and Teaching in the Social Sciences, University of Colorado, Boulder, CO

STEM RRG Project: http://www.tamiu.edu/~rbachnak/STEMRRG/index3.html

Appendix

STUDENT SELF EVALUATION OF PRESENTATION

POSTER/ ORAL Presentation: _____

Graduate/ Undergraduate: _____

Major/ Minor: _____

STEM discipline: Math/Bio/Stat/CS: _____

Senior/Junior/Sophomore/Freshman: _____

Evaluate your presentation on _____@_____ on a scale of 1—5, 5 being the superior for each of the following. (For Example: 1 = inadequate, 2 = mediocre, 3 = satisfactory, 4 = good, 5 = very good).

1) How do you rate your overall presentation?
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

2) What is the extent of learning outside the classroom activities?
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

3) What was the extent of your thoroughness of the topic in your presentation?
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

4) What was the extent of difficulty of getting this presentation done in time?
   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

5) Did you rehearse your presentation before actually presenting?
   ☐ Yes ☐ No

6) Did you seek outside help?
☐ Yes ☐ No

7) What is the amount of time spent preparing for the project?
☐ 1 day ☐ 1 week ☐ 1 month ☐ 2 months

8) When did you start gathering information on the topic?
☐ 1 week ago ☐ 2 weeks ago ☐ 3 weeks ago ☐ 4 weeks ago

9) How many sources did you consult?
☐ 0 ☐ 0-4 ☐ 4-8 ☐ 8-12 ☐ 12 or more

10) Is the topic course-related/ project-related?
☐ Course ☐ Project

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 2, 24 – 26, 2009

A NEW METHOD FOR SOLVING QUARTICS

RAGHAVENDRA G. KULKARNI
Senior Member IEEE, Life Member Ramanujan Mathematical Society
Senior Deputy General Manager, HMC division
Bharat Electronics Ltd.,
Jalahalli Post, Bangalore-560013, INDIA.
Phone: +91-80-22195270, Fax: +91-80-28382738, Email: rgkulkarni@ieee.org
http://www.bel-india.com

A new method for solving the general quartic equation is presented, which is different from the methods available in the literature.

Keywords: Quartic equation, cubic equation, polynomial decomposition

1. Introduction

In this paper we describe a new method to solve the general quartic equation. The method is different from the well-known Ferrari's method, or any other earlier method [Wikipedia, Kulkarni (2006)]. In this method the given general quartic equation is compared with an easily decomposable quartic equation, resulting in its decomposition into a product of two quadratic polynomials.

2. The proposed method

Consider the general quartic equation in x, for which solution is sought, as given below.

$$x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0 \quad (1)$$

where the coefficients, a_0, a_1, a_2, and a_3, are real. Consider another quartic equation shown below.

$$(x^2 + b_1x + b_0)^2 - p(x + c_0)^2 = 0 \quad (2)$$

where b_0, b_1, c_0, and p are unknowns to be determined. We attempt to represent the given quartic (1) in the form of (2) which is easily decomposable into two factors as indicated below.

$$[x^2 + (b_1 - p^{1/2})x + b_0 - p^{1/2}c_0][x^2 + (b_1 + p^{1/2})x + b_0 + p^{1/2}c_0] = 0 \quad (3)$$

For this purpose, we expand (2) and rearrange as shown below to facilitate comparison with (1).

$$x^4 + 2b_1x^3 + (b_1^2 + 2b_0 - p)x^2 + 2(b_0b_1 - pc_0)x + b_0^2 - pc_0^2 = 0 \quad (4)$$

Equating the corresponding coefficients of quartic equations (1) and (4), we obtain four equations in four unknowns, b_0, b_1, c_0, and p, as shown below.

$$2b_1 = a_3 \quad (5)$$

$$b_1^2 + 2b_0 - p = a_2 \quad (6)$$

$$2(b_0b_1 - pc_0) = a_1 \quad (7)$$

$$b_0^2 - pc_0^2 = a_0 \quad (8)$$

Using the above four equations the unknowns, b_0, b_1, c_0, and p, are determined. From (5) b_1 is evaluated as, $b_1 = a_3/2$, and from (5) and (6) b_0 is expressed as shown below.

$$b_0 = [p + a_2 - (a_3^2/4)]/2 \quad (9)$$

Using (9) we eliminate b_0 from equation (7), to get an expression in p and c_0, as shown below.

$$2pc_0 = (a_3/2)[p + a_2 - (a_3^2/4)] - a_1 \tag{10}$$

Again using (9), b_0 is eliminated from (8) resulting in another expression in p and c_0 as given below.

$$4pc_0^2 = [p + a_2 - (a_3^2/4)]^2 - 4a_0 \tag{11}$$

Squaring equation (10) and then dividing it by (11), we eliminate c_0, and obtain an expression in p as shown below.

$$p\{[p + a_2 - (a_3^2/4)]^2 - 4a_0\} = \{(a_3/2)[p + a_2 - (a_3^2/4)] - a_1\}^2 \tag{12}$$

To simplify equation (12) we make following substitutions.

$$F_1 = a_2 - (a_3^2/4)$$
$$p_1 = p + F_1 \tag{13}$$

With these substitutions and after some rearrangement, equation (12) will appear as:

$$p_1^3 - a_2 p_1^2 + (a_1 a_2 - 4a_0)p_1 + 4a_0 F_1 - a_1^2 = 0 \tag{14}$$

Notice that (14) is a cubic equation in p_1 with real coefficients, and when solved over the real field it provides at least one real root. Once p_1 is known p is evaluated from (13), and then, b_0 and c_0 are found out from (9) and (10) respectively. Having determined all the unknowns, we are in a position to represent the given quartic equation (1) in the form of (2), whose factors are quadratic polynomials as given in (3). Equating each of these factors to zero, and solving the resulting quadratic equations, we get all the four roots of given quartic (1) as shown below.

$$x_1 = \{ -(b_1 - p^{1/2}) + [(b_1 - p^{1/2})^2 - 4(b_0 - p^{1/2}c_0)]^{1/2}\}/2$$
$$x_2 = \{ -(b_1 - p^{1/2}) - [(b_1 - p^{1/2})^2 - 4(b_0 - p^{1/2}c_0)]^{1/2}\}/2$$
$$x_3 = \{ -(b_1 + p^{1/2}) + [(b_1 + p^{1/2})^2 - 4(b_0 + p^{1/2}c_0)]^{1/2}\}/2$$
$$x_4 = \{ -(b_1 + p^{1/2}) - [(b_1 + p^{1/2})^2 - 4(b_0 + p^{1/2}c_0)]^{1/2}\}/2 \tag{15}$$

3. Numerical example

Let us solve the following quartic equation using the method proposed here:

$$x^4 - x^3 - 19x^2 - 11x + 30 = 0.$$

For this quartic equation b_1 is found out as: $b_1 = -1/2$. F_1 is evaluated as: $F_1 = -(77/4)$, and the cubic equation (14) in p_1 is given by:

$$p_1^3 + 19p_1^2 - 109p_1 - 2431 = 0.$$

The above cubic is solved [Kulkarni, 2006], and its roots are obtained as: 11, – 13, – 17. With $p_1 = 11$, we determine p, b_0, and c_0 as: $p = 121/4$, $b_0 = 11/2$, and $c_0 = 1/11$. We calculate back the coefficients, a_0, a_1, a_2 and a_3, and obtain the values as: 30, – 11, – 19, and – 1, respectively. Thus we ensure that, $p_1 = 11$, is the desired value required to be used to find the roots of quartic in this example. Using the expressions given in (15), the roots of quartic, x_1, x_2, x_3 and x_4, are determined as: 5, 1, – 2, and, – 3 respectively.

4. Conclusions

The general quartic equation is solved using a new method, which compares the given quartic with an easily decomposable quartic. A numerical example is solved using the proposed method.

Acknowledgments

The author thanks the management of Bharat Electronics Ltd., Bangalore for supporting this work.

References

http://en.wikipedia.org/wiki/Quartic_equation.
Kulkarni, R. G. (2006): "Unified method for solving general polynomial equations of degree less than five" Alabama Journal of Mathematics, **30**(1 & 2), pp. 1 – 18.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 2, No. 2, 27 – 30, 2009

A NUMBER THEORETIC IDENTITY

Amiya Bhushan Sarkar
Department of Mathematics,
Goa University, Taleigao Plateau,
Goa-403206,[1] Mobile: +8801720111727
sarkaramiya737@gmail.com

In this note we discuss a certain number theoretic identity with some interesting connections. It is seen that not only can this identity be extended to more general situations, but it also has close connections with the notion of a congruent number. In fact the above-mentioned identity holds good in the case of square matrices as well as bounded linear operators. Beside some interesting geometric results emerge naturally as a consequence of this identity.

Keywords. Number; theoretic; identity; congruent; triple; pythagoras.

1. Introduction

The number theoretic identity we are interested in is a generalization of the identity, $x^2 + (x+1)^2 + (x(x+1))^2 = [x(x+1)+1]^2$ (1)

If we expand both the sides separately then it is seen that the equality clearly holds. However what is interesting is the fact that the above equality arises as a special case of the identity,

$x^2 + (x+1)^2 + (p+m)^2 = (q+m)^2 \pm 2m$, where $p = x(x+1)$, $q = x(x+1) + 1$ and $m = 0, \pm 1, \pm 2$,.....

i) $x^2 + (x+1)^2 + (p+m)^2 = (q+m)^2 - 2m$, when m = 0, 1, 2.....

ii) $x^2 + (x+1)^2 + (p-m)^2 = (q-m)^2 + 2m$, when m = 0, -1, -2.....

It is seen that by simply setting m = 0, we get identity (1).

Now by expanding both the left hand side and the right hand side separately, the validity of this identity can be easily established for all $x \in R$.

For the case m = 0, we have a generalization of the identity (1). To this end we have the following theorem:

Theorem 1.1. *For* $x \in R$, $\sum_{k=0}^{2n+1}(x+k)^2 = 2[x(x+1)+.....+(x+2n)(x+2n+1)] + (n+1)$

for n= 0, 1, 2, 3.,.....

Proof: This can be established by appealing to the principle of mathematical induction.

2. Extension to the Matrix and the Bounded Linear Operator Cases

It is interesting to note that this identity holds good in the case of both square matrices as well as bounded linear operations on a linear space X. To this end we have the following proposition.

Proposition 2.1 *If A is a n x n square matrix or a bounded linear operator on a linear space X, then ,*

(I) $A^2 + (A+I)^2 + [A(A+I) + mI]^2 = [A(A+I) + I + mI]^2 - 2mI,$

(II) $A^2 + (A+I)^2 + [A(A+I) - mI]^2 = [A(A+I) + I - mI]^2 + 2mI,$

where I is the identity matrix or the identity linear operator, as the case may be, and m = 0, ± 1, ± 2, ...

Proof. The proof is along the same lines as in the case of the number theoretic identity. On expanding both sides of the equality separately we get the desired result.

It is worth noting that the above proposition is actually a consequence of a more general result stated below.

Theorem 2.2. *If A and B are two n x n matrices (or bounded linear operators on a linear space X), such that AB = BA and* $B^2 = B$*, than,*

$(AB)^2 + (AB+B)^2 + [A(A+B)]^2 = [A(A+B)+B]^2.$

Proof. The validity of this equality is established by expanding both the left hand side and the right hand side separately.

It is not difficult to see that by setting B = 1 in the theorem we recover proposition 2.1, for m = 0

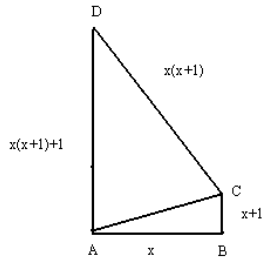
3. Some Geometry

Consider a right angled triangle ABC with the right angle subtended at the vertex B as shown in the figure below, with AB and BC having lengths x and x+1 respectively. Next we construct another angled triangle ACD with DC perpendicular to AC, and DC having length x(x+1). Now it follows from Pythagoras' theorem that,

$AB^2 + BC^2 + DC^2 = AD^2$

i.e, $x^2 + (x+1)^2 + [x(x+1)]^2 = AD^2$

It follows from the identity established earlier that $AD^2 = [x(x+1)+1]^2$.



Given this scenario we immediately see a close connection with the notion of a congruent number.

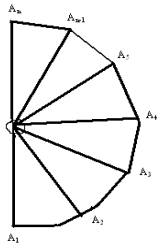
Definition 3.1. *A positive integer n is said to be a congruent number if there exist v, w, x, y, z rational numbers, such that* $v^2 + w^2 + x^2 + y^2 = z^2$ *and* $\frac{vw}{2} + \frac{xy}{2} = n$.

If in the discussion preceding this definition x is an integer, then it follows from the equality $x^2 + (x+1)^2 + (\sqrt{(x^2+(x+1)^2)})^2 + \{x(x+1)\}^2 = [x(x+1)+1]^2 + [x^2+(x+1)]^2$, that

$n = \frac{x(x+1)}{2} + \frac{\sqrt{(x^2+(x+1)^2)(x(x+1))}}{2}$ is a congruent number so long as $(x^2 + (x+1)^2)$ is a perfect square. In this case it turns out that n is the sum of the areas of the two triangles.

Next, we prove the following theorem in an attempt to generalize the notion of Pythagorean Triple [2].

Theorem 3.2. *Consider the polygon shown below, comprising of n right angled triangles with every pair of adjacent triangles having a common side. We claim that* $OA_n^2 = OA_1^2 + A_1A_2^2 + A_2A_3^2 + \ldots\ldots + A_{n-1}A_n^2$



Proof. The proof follows from the Pythagoras theorem and the principle of mathematical induction.

4. Conclusion

In this paper we have discussed an interesting number theoretic identity, which has several generalizations and is closely connected with the notion of a congruent number. Here we have generalized this notion through a geometric approach.

Acknowledgement

It is a pleasure to think Dr. E. De Sa, Dr. A. Mohopatra, Mr. Kare and Mrs. Sheryl Dias Barreto for their support and encouragement. Of course there are others who also deserve special mention especially those who have always stood by me. My sincere thanks to all of them.

References

(1) Chandrashekar V, The congruent number problems, Mathematics Newsletter, 2 (1998) 17-22.

(2) Venkataraman J, Pythagorean triples, The Mathematics Teacher. 32 (1996) 36-41.

ISSN 0974 – 3340

SUTRA

International Journal of

Mathematical Science Education

VOLUME 3, ISSUE 1 JANUARY-JUNE 2010

Published by

Technomathematics Research Foundation, India

Sutra: International Journal of Mathematical Science Education

EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India.*

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain.*
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK.*
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India.*
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India.*
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India.*
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India.*
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India.*
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia.*
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

Copyright ©2010 *Technomthematics Research Foundation.*

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.

Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.

Sutra: International Journal of Mathematical Science Education (ISSN 0974 – 3340) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education

CONTENTS (Vol. 3, No. 1, 2010)

1. MEASURE OF PLANES INTERSECTING A CONVEX BODY — 1—7
   Rafik Aramyan

2. THE SUM OF ONE SERIES — 8—15
   Euler and Bernoulli summation methods of inverted squares series
   Frantisek Mosna

3. BOUNDED INVERSE RELATIONAL FUNCTION — 16—23
   Chandan Kumar

4. NUMBERS : RATIONAL, IRRATIONAL OR TRANSCENDENTAL ? — 24—30
   Kallol Paul and Dipanwita Paul Ghosh

5. A MIXED QUADRATURE RULE FOR NUMERICAL INTEGRATION OF ANALYTIC FUNCTIONS BY USING BIRKHOFF YOUNG AND BOOLS QUADRATURE — 31—35
   Sanjit Ku. Mohanty and Rajani.B. Dash

6. ETHNOMATHEMATICS: A Mathematical Game in Hausa Culture — 36—42
   Mohammed Waziri Yusuf, Ibrahim Saidu, and *Aisha Halliru*

7. A COEFFICIENT INEQUALITY FOR A SUB-CLASS OF ANALYTIC FUNCTIONS — 43—52
   B. S. Mehrok, Deepak Gupta and Gagandeep Singh

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3 No. 1, pp. 1 - 7, 2010

MEASURE OF PLANES INTERSECTING A CONVEX BODY

RAFIK ARAMYAN

Russian-Armenian (Slavonic) University,
Institute of Mathematics Armenian Academy of Sciences,
Bagramian 24b, Yerevan, ARMENIA
rafikaramyan@yahoo.com

In the present paper a representation for the measure of planes intersecting a convex body, using stochastic approximation of the body was found. The representation was found in terms of the normal curvatures of the surface of the body and a flag density of the measure.

Keywords: integral geometry; stochastic approximation; flag densities.

1. Introduction

Let $\mathbf{E}$ be the space of planes in $\mathbf{R}^3$. We consider locally finite signed measures μ in the space $\mathbf{E}$, which possess densities with respect to the standard Euclidean motion invariant measure,
i. e. (see. [**?**])

$$\mu(de) = h(e)\, de. \tag{1}$$

Recall that an element de of the standard measure is written as $de = dp \cdot d\xi$, where (p, ξ) is the usual parametrization of a plane e: p is the distance of e from the origin O; $\xi \in \mathrm{S}^2$ is the direction normal to e, $d\xi$ is an element of solid angle of the unit sphere S^2. Where appropriate we write $h(e) = h(p, \xi)$.

The concept of a flag in $\mathbf{R}^3$ which naturally emerges in Combinatorial integral geometry will be of basic importance below. A detailed account of this concept is in [**?**]. We repeat the definition.

A flag is a triad $f = (P, g, e)$, where P is a point in $\mathbf{R}^3$ called the *l*ocation of f, g is a line containing the point P, and e is a plane containing g. There are two equivalent representations of a flag:

$$f = f(P, \Omega, \Phi) \text{ or } f = f(P, \omega, \varphi),$$

where Ω is the spatial direction of g in $\mathbf{R}^3$, Φ is the rotation of e around g, ω is the normal of e, and φ is the planar direction of g in e. The range of Ω and ω is

$\mathcal{E}_2$, the standard elliptic 2–space which can be obtained from the unit sphere by identification of the antipodal points ([**?**]), ϕ and φ belong to $\mathcal{E}_1$.

We introduce the following function in the space of flags (a flag function)

$$\rho(f) = \rho(P, \omega, \varphi) = \int_{\mathcal{E}_2} \cos^2(\varphi - \psi)\, h_{[P]}(\xi)\, d\xi. \tag{2}$$

Here $[P]$ is the bundle of planes containing the point $P \in \mathbf{R}^3$, $h_{[P]}(\xi)$ is the restriction of h onto $[P]$, ψ is the direction of the projection of ξ into the plane of the flag f. The notation $h_{[P]}(\xi)$ is reasonable since ξ completely determines a plane from $[P]$. Clearly, the integral (2) does not depend on the choice of the reference point on the plane of the flag f. The function ρ defined on the space of flags $\mathcal{F}$ we call flag density. The concept of a flag density was introduced and systematically employed by R. V. Ambartzumian (see [**?**],[**?**]).

Note, that in [**?**] (see also [**?**] and [**?**]) (2) was considered as an integral equation and by integral geometry methods was recovered h from a given ρ.

Let $\mathbf{B}$ be a convex body with sufficiently smooth boundary and with positive Gaussian curvature at every point of $\partial\mathbf{B}$. By $[\mathbf{B}]$ we denote the set of planes intersecting $\mathbf{B}$. Let $s(\omega)$ be a point on $\partial\mathbf{B}$ whose outer normal is ω. By $k_1(\omega), k_2(\omega)$ we denote the principal normal curvatures of $\partial\mathbf{B}$ at $s(\omega)$ and let $k(\omega, \varphi)$ be the normal curvature in the direction φ at the point $s(\omega)$ of $\partial\mathbf{B}$, φ is measured from the first principal direction.

The main result of the paper is the following.

Theorem 1.1. *Let μ be a signed measure on $\mathbf{E}$, possessing a density $h(e)$. For any sufficiently smooth convex body $\mathbf{B}$ we have the following representation:*

$$\mu([\mathbf{B}]) = (2\pi^2)^{-1} \int_{\mathrm{S}^2} \int_0^{2\pi} \rho(s(\omega), \omega, \varphi) \frac{\sqrt{k_1(\omega) k_2(\omega)}}{k^2(\omega, \varphi)} d\varphi\, d\omega, \tag{3}$$

where ρ is the flag density of μ defined by (2).

If $h(e) \equiv 1$ (the case of Euclidean motion invariant measure μ_{inv}) from (3) we obtain the Minkowski formula (see [**?**])

$$\mu_{\mathrm{inv}}([\mathbf{B}]) = \frac{1}{2} \int_{\mathrm{S}^2} \left(\frac{1}{k_1(\omega)} + \frac{1}{k_2(\omega)} \right) d\omega. \tag{4}$$

2. Preliminary Representations

In [**?**] R. V. Ambartzumian has indicated the existence of the so-called *flag-representation* for width functions of convex bodies in $\mathbf{R}^3$ using some "standard" flag-representation for width functions of polyhedra.
Let $H(\xi)$ be the width function in direction ξ of a convex body $\mathbf{B}$. Then (see [**?**])

$$H(\xi) = \int_{\mathrm{S}^1 \times \mathrm{S}^2} \sin^2 \alpha(\xi, \Omega, \Phi)\, m(d\Omega, d\Phi), \tag{5}$$

where S^i is the unit sphere in $\mathbf{R}^{i+1}$, $i = 1, 2$, Ω, $\xi \in \mathrm{S}^2$, m is a measure in $\mathrm{S}^1 \times \mathrm{S}^2$, α is the angle between $\Omega \in \mathrm{S}^2$ and the trace $e_\xi \cap e(\Omega, \Phi)$, e_ξ is a plane normal to ξ, and $e(\Omega, \Phi)$ is the plane of the so-called free flag $f = f(\Omega, \Phi)$ (one can consider that the location of a free flag is the origin of $\mathbf{R}^3$).
The representation (5) fails to be unique (there are many m for given H).
Note, that if μ is a translation invariant measure in $\mathbf{E}$ with the form $d\mu = dp \times \delta_\xi$, where δ_ξ is a delta measure concentrated on the direction ξ, then we have $H(\xi) = \mu([\mathbf{B}])$.

Let $\mathbf{K} \subset \mathbf{R}^3$ be a convex polyhedron and $e \in [\mathbf{K}] \subset \mathbf{E}$. We consider the intersection $e \cap \mathbf{K}$ which is a bounded convex polygon whose vertices correspond to the edges of $\mathbf{K}$ actually hit by e. The fact, that the sum of outer angles of $e \cap \mathbf{K}$ equals 2π we write in the form

$$\sum_i \alpha_i(e)\, I_{[L_i]}(e) = 2\pi I_{[\mathbf{K}]}. \tag{6}$$

Here L_i is an edge of $\mathbf{K}$, $\alpha_i(e)$ is the outer angle of $e \cap \mathbf{K}$ correspond to vertex $e \cap L_i$, and summation is by all edges of $\mathbf{K}$.
In [**?**], by integration of (6) with respect to $d\mu = dp \times m(d\xi)$ (a translation invariant measure) some "standard" flag-representation for the width function of a polyhedron $\mathbf{K}$ was found. In [**?**], using approximation by polyhedrons, a new representation for the width functions of convex bodies was obtained. In [**?**], using stochastic approximation (Voronoi's approximation) of smooth convex bodies by polyhedrons, for translation invariant measures representation (3) was obtained.
Now we integrate (6) with respect to $\mu(de) = h(e)de$, where h is a continuous function defined on $\mathbf{E}$. We have

$$2\pi\mu([\mathbf{K}]) = \sum_i \int_{[L_i]} \alpha_i(e)\, h(e)\, de = \sum_i \int_{[L_i]} \alpha_i(e)\, h(p, \xi)\, dp\, d\xi$$

$$= \sum_i \int_{[L_i]} \alpha_i(e) h(x, \xi) \mid \cos(\widehat{\xi, \Omega_i}) \mid dx d\xi = \sum_i \int_{L_i} \int_{\mathcal{E}_2} \alpha_i(e) h(x, \xi) \mid \cos(\widehat{\xi, \Omega_i}) \mid d\xi dx. \tag{7}$$

Here $\widehat{\xi, \Omega_i}$ is the angle between ξ and Ω_i, where Ω_i is the direction of the edge L_i. Also, here we use the following well known fact from integral geometry

$$de = dp\, d\xi = \mid \cos(\widehat{\xi, \Omega_i}) \mid \; dx\, d\xi, \tag{8}$$

where x is the intersection point $e \cap L_i$ and dx is one dimensional Lebesque in L_i. Using standard formulae of spherical trigonometry we get (see [**?**])

$$\alpha_i(e) \; \mid \cos(\widehat{\xi, \Omega_i}) \mid = \int_{A_i} \sin^2 \alpha(\xi, \Omega_i, \Phi)\, d\Phi, \tag{9}$$

where A_i is the exterior dihedral angle of the edge L_i (see also (5)). After substitution (9) into (7) we obtain

$$2\pi\mu([\mathbf{K}]) = \sum_i \int_{L_i} \int_{A_i} \int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_i, \Phi)\, h(x, \xi) d\xi\, d\Phi\, dx. \tag{10}$$

3. Stochastic Approximation

Let $\mathbf{B}$ be a sufficiently smooth (three times continuously differentiable) convex body in $\mathbf{R}^3$. We assume that the Gaussian curvature of $\partial\mathbf{B}$ is everywhere positive. Hence the Gauss map of $\partial\mathbf{B}$ onto the unit sphere S^2 is a homeomorphism.
We throw n independent points $P_1,\dots,P_n$ onto S^2 with the same distribution P. Let $d\mathrm{P}=f(\omega)d\omega$, where $f(\omega)>0$ is continuous, $d\omega$ is an area element on S^2. On $\partial\mathbf{B}$ by $P_1^*,\dots,P_n^*$ we denote the images of the points $P_1,\dots,P_n$ by the inverse to the Gauss map. Denote by $\mathbf{K_n}(P_1^*,\dots,P_n^*)$ the convex hull of the points $P_1^*,\dots,P_n^*$. According to (10), $\mu([\mathbf{K_n}(P_1^*,\dots,P_n^*)])$ can be represented in the form

$$2\pi\mu([\mathbf{K_n}])=\sum_{i<j\ \ i,j=1}^{n} I_D(i,j)\int_{L_{ij}}\int_{A_{ij}}\int_{\mathcal{E}^2}\sin^2\alpha(\xi,\Omega_{ij},\Phi)\,h(x,\xi)d\xi\,d\Phi\,dx. \quad (11)$$

Here Ω_{ij} is the direction of $\overrightarrow{P_i^*P_j^*}$, L_{ij} is the edge $P_i^*P_j^*$, A_{ij} is the exterior dihedral angle of the edge $P_i^*P_j^*$, D is the set of all pairs (i,j) corresponding to the edge. We average both sides of (11) with respect to the sequences $(P_1^*,\dots,P_n^*)$. Since $f(\omega)>0$, in the limit $(n\to\infty)$ in the left-hand side we obtain $\mu([\mathbf{B}])$. By symmetry we have

$$2\pi\mu([\mathbf{B}])=\lim_{n\to\infty}\binom{n}{2}\int_{(\mathrm{S}^2)^2}\left[\int_{(\mathrm{S}^2)^{n-2}} I_D(1,2)\right.$$
$$\left.\times\left[\int_{L_{12}}\int_{A_{12}}\int_{\mathcal{E}_2}\sin^2\alpha(\xi,\Omega_{12},\Phi)\,h(x,\xi)d\xi\,d\Phi\,dx\right]dP_3\cdots dP_n\right]dP_1\,dP_2. \quad (12)$$

Taking P_1 as the pole, P_2 can be described by spherical coordinates (ν,φ) with respect to P_1. Also, P_2 can be described by coordinates (l,φ), where $l=|P_1P_2|$. We have

$$dP_2=f(\omega)\,d\omega=f(\nu,\varphi)\,\sin\nu\,d\nu\,d\varphi=f(l,\varphi)\,l\,dl\,d\varphi. \quad (13)$$

Let $e(\Omega_{l\varphi},\Phi)$ be the plane passing through P_1^*,P_2^* and rotated around $\Omega_{l\varphi}=\overrightarrow{P_1^*P_2^*}$ by angle Φ. For $e(\Omega_{l\varphi},0)$ we take the plane that is perpendicular to the plane passing through ω and $\Omega_{l\varphi}$. By L^* we denote the segment $P_1^*P_2^*$ and let $l^*=|P_1^*P_2^*|$.
In this paper we consider the case of the uniform distribution, i.e. $f(\omega)=(S_oC(\omega))^{-1}$, where S_o is the total area of the surface of $\mathbf{B}$, $C(\omega)$ is the Gaussian curvature at the point on $\partial\mathbf{B}$ with normal ω. The plane $e(\Omega_{l\varphi},\Phi)$ divides $\partial\mathbf{B}$ into two parts and by $S(\Phi,l)$ we denote the area of the smaller part $\partial\mathbf{B}_1(\Phi,l)$. Applying Fubini's theorem in the inner integral of (12), we obtain

$$2\pi\mu([\mathbf{B}])=\lim_{n\to\infty}\binom{n}{2}\int_{(\mathrm{S}^2)^2}\left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\left[\left(1-\frac{S(\Phi,l)}{S_o}\right)^{n-2}+\left(\frac{S(\Phi,l)}{S_o}\right)^{n-2}\right]\right.$$
$$\left.\times\left[\int_{L^*}\int_{\mathcal{E}_2}\sin^2\alpha(\xi,\Omega_{l\varphi},\Phi)\,h(x,\xi)d\xi\,dx\right]d\Phi\right]\frac{l\,dl\,d\varphi\,d\omega}{S_o^2C(\omega)C(l,\varphi)}. \quad (14)$$

The sum in the square brackets of (14) is the probability that segment $P_1^*P_2^*$ is an edge and $e(\Omega_{l\varphi},\Phi)$ belongs to the exterior dihedral angle of the edge.

Since $\frac{S(\Phi,l)}{S_o} \leq \frac{1}{2}$, we have

$$\lim_{n\to\infty} | \binom{n}{2} \int_{(\mathrm{S}^2)^2} \left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{S(\Phi,l)}{S_o} \right)^{n-2} \left[\int_{L^*} \int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_{l\varphi}, \Phi)\, h(x,\xi) d\xi\, dx \right] d\Phi \right] \times \frac{l\, dl\, d\varphi\, d\omega}{S_o^2 C(\omega) C(l,\varphi)} | \leq \lim_{n\to\infty} A \binom{n}{2} \left(\frac{1}{2}\right)^{n-2} = 0, \tag{15}$$

where A is a constant. In a similar manner one can prove that the domain of variation of Φ and l can be taken arbitrarily small. Thus

$$2\pi\mu([\mathbf{B}]) = \lim_{n\to\infty} \binom{n}{2} \int_{(\mathrm{S}^2)} \int_0^{2\pi} \int_0^{l_0} \int_{-\Phi_0}^{\Phi_0} \left(1 - \frac{S(\Phi,l)}{S_o}\right)^{n-2} \times \left[\int_{L^*} \int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_{l\varphi}, \Phi)\, h(x,\xi) d\xi\, dx \right] l\, d\Phi\, dl \frac{d\varphi\, d\omega}{S_o^2 C(\omega) C(l,\varphi)}, \tag{16}$$

where l_0 and Φ_0 are arbitrarily small fixed numbers. From the regularity of the surface $\partial\mathbf{B}$ we obtain the Taylor expansion

$$S(\Phi,l) = l\, S'_l(0,0) + \Phi S'_\Phi(0,0) + \frac{l^2}{2} S''_{ll}(0,0) + l\Phi S''_{l\Phi}(0,0) + \frac{\Phi^2}{2} S''_{\Phi\Phi}(0,0) + R(\Phi,l), \tag{17}$$

where $R(\Phi,l) = o(l^2+\Phi^2)$. Here all functions continuously depend on l and Φ, as well as on ω and φ. Below, we will see, that $S'_l(0,0) = S'_\Phi(0,0) = 0$.
Using the mean value theorem we find

$$\int_{L^*} \int_{\mathcal{E}^2} \sin^2 \alpha(\xi, \Omega_{l\varphi}, \Phi)\, h(x,\xi) d\xi\, dx = l^* \int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_{l\varphi}, \Phi)\, h(x_0,\xi) d\xi, \tag{18}$$

where x_0 is a point from the segment L^* and $l^* = |L^*|$.
After substitution (18) in (16) and a change of variables $u = l\sqrt{n}$, $v = \Phi\sqrt{n}$ we get

$$2\pi\mu([\mathbf{B}]) = \lim_{n\to\infty} \binom{n}{2} \int_{(\mathrm{S}^2)} \int_0^{2\pi} \left[\int_0^{l_0\sqrt{n}} \int_{-\Phi_0\sqrt{n}}^{\Phi_0\sqrt{n}} \left(1 - \frac{S(\frac{v}{\sqrt{n}}, \frac{u}{\sqrt{n}})}{S_o}\right)^{n-2} \times \left[\int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_{\frac{u}{\sqrt{n}}\varphi}, \frac{v}{\sqrt{n}})\, h(x_0,\xi) d\xi \right] \frac{u(u\cdot b(\omega,\varphi)) du\, dv}{n^2 C(\frac{v}{\sqrt{n}}, \varphi)} \right] \frac{d\varphi\, d\omega}{S_o^2 C(\omega)}, \tag{19}$$

where $l^* = l \cdot b(\omega,\varphi) + o(l)$.
One can interchange the limit and the integration operations. Substitution (17) in (19) and making use of

$$\int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \Omega_{\frac{u}{\sqrt{n}}\varphi}, \frac{v}{\sqrt{n}})\, h(x_0,\xi) d\xi \to \int_{\mathcal{E}_2} \sin^2 \alpha(\xi, \omega, \varphi_1)\, h_{[P^*(\omega)]}(\xi) d\xi \tag{20}$$

almost everywhere when $n \to \infty$, where $\alpha(\xi,\omega,\varphi_1)$ is the angle between the direction φ_1 in the plane e_ω and the intersection of e_ξ with the plane e_ω, $[P^*(\omega)]$ is the

bundle of planes containing the point $P^*(\omega) \in \partial\mathbf{B}$ with normal ω, we obtain

$$2\pi\mu([\mathbf{B}]) = \int_{(\mathrm{S}^2)} \int_0^{2\pi} \left[\int_0^\infty \int_{-\infty}^\infty \exp\left(-\frac{u^2}{2} S''_{ll}(0,0) - uv S''_{l\Phi}(0,0) - \frac{v^2}{2} S''_{\Phi\Phi}(0,0) \right) \right.$$
$$\left. \times\, u^2 du\, dv \right] \left[\int_{\mathcal{E}_2} \sin^2 \alpha(\xi,\omega,\varphi_1)\, h_{[P^*(\omega)]}(\xi) d\xi \right] \frac{b(\omega,\varphi)}{2C(\omega)} d\varphi\, d\omega, \tag{21}$$

where $b(\omega,\varphi) = \lim_{l\to 0} \frac{l^*}{l}$.

4. A Representation Obtained by Stochastic Approximation

It follows from (21) that the final representation for $\mu([\mathbf{B}])$ depends on values of $S''_{ll}(0,0)$, $S''_{l\Phi}(0,0)$, $S''_{\Phi\Phi}(0,0)$ which are functions of ω and φ (see (21)). It was proved in [**?**], that:

$S'_l(0,0)$, $S'_\Phi(0,0)$, $S''_{ll}(0,0)$, $S''_{l\Phi}(0,0)$, $S''_{\Phi\Phi}(0,0)$ depend only on derivatives of at most order of two of the surface $\partial\mathbf{B}$ at the point P^* whose outer normal is ω.
Hence the corresponding calculation we can do for the osculating paraboloid of $\partial\mathbf{B}$ at the point $P^*(\omega)$ whose outer normal is ω. In [**?**] the following expressions for the derivatives in terms of the normal curvatures of $\partial\mathbf{B}$ at the point $P^*(\omega)$ were found:

$$S'_l(0,0) = 0, \quad S'_\Phi(0,0) = 0, \quad S''_{ll}(0,0) = \frac{\pi\sqrt{k_1 k_2}\, r^2(\varphi)(k_2^3 \cos^2\varphi + k_1^3 \sin^2\varphi)}{2A^4},$$
$$S''_{l\Phi}(0,0) = \frac{\pi\sqrt{k_1 k_2}\sin 2\varphi(k_2 - k_1)}{2A^3}, \quad S''_{\Phi\Phi}(0,0) = \frac{2\pi\sqrt{k_1 k_2}\, r(\varphi)}{A^2}, \tag{22}$$

where k_i, $i = 1,2$ are the main normal curvatures, $r(\varphi) = k_1^{-1}\cos^2\varphi + k_2^{-1}\sin^2\varphi$ is the radius of the normal curvature in the direction φ at the point $P^*(\omega)$ of $\partial\mathbf{B}$ and $A = \sqrt{k_2^2\cos^2\varphi + k_1^2\sin^2\varphi}$. Also, in [**?**] was found that (see (21))

$$b(\omega,\varphi) = \sqrt{\frac{\cos^2\varphi}{k_1^2} + \frac{\sin^2\varphi}{k_2^2}} \quad \text{and} \quad \tan\varphi_1 = \tan\varphi\, \frac{k_1}{k_2}. \tag{23}$$

Substituting (22) and (23) into (21) we get

$$\mu([\mathbf{B}]) = (2\pi^2)^{-1} \int_{(\mathrm{S}^2)} \int_0^{2\pi} \left[\int_{\mathcal{E}^2} \sin^2 \alpha(\xi,\omega,\varphi)\, h_{[P^*(\omega)]}(\xi) d\xi \right] \frac{\sqrt{k_1 k_2}}{k^2(\omega,\varphi)} d\varphi\, d\omega, \tag{24}$$

where $k(\omega,\varphi)$ is the normal curvature in the direction φ at the point $P^*(\omega)$ of $\partial\mathbf{B}$. Taking into account that

$$\sin^2\alpha(\xi,\omega,\varphi) = \cos^2(\varphi - \psi), \tag{25}$$

where ψ is the direction of the projection of ξ into the plane with normal ω, we get (3).

Theorem 1 is proved.

I would like to express my gratitude to Prof. R. V. Ambartzumian for helpful discussions.

References

Ambartzumian, R. V. (1987). Combinatorial integral geometry, metric and zonoids. *Acta Appl. Math.*, **9**: 3-27.

Ambartzumian, R. V. (1990). *Factorization Calculus and Geometric Probability)*, Cambridge University Press, Cambridge.

Aramyan, R. G. (1986). About stochastic approximation of convex bodies. *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)*, **22**: 427-438.

Aramyan, R. G. (1992). Recovering rose of directions from flag densities in $\mathbf{R}^3$. *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)*, **27**: 421-429.

Aramyan, R. G. (2009). Solution of one integral equation on the sphere by methods of integral geometry. *Doklady Mathematics*, **79**: 325-328.

Panina, G. Yu. (1986). Convex bodies and translation invariant measures. *Zap. Nauchn. Semin. LOMI*, **157**: 143-152.

Ambartzumian, R. V., Ohanyan, V. K. (1994). Finite additive functionals in the space of planes. *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)*, **29**: 3-51.

Santalo, Luis A. (1976). *Integral Geometry and Geometric Probability*, Addison-Wesley Publishing Company, Canada.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3 No. 1, pp. 8 - 15, 2010

THE SUM OF ONE SERIES
Euler and Bernoulli summation methods of inverted squares series

FRANTIŠEK MOŠNA*

Department of Mathematics, Czech University of Life Sciences, Kamýcká 129
16531 Praha 6, Czech Rep.,
mosna@tf.czu.cz
http://mosna.tf.czu.cz

The series of inverted squares and its sum can be investigated in several ways. We offer two of them - the Euler method and Bernoulli method using product of series.

Keywords: Infinite series; Euler; Bernoulli; product of series.

The series $\sum \frac{1}{n^2}$ is convergent and its sum is

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \ .$$

We usually prove it using e.g. Fourier series and the Parseval equality. The purpose of this text is to recollect another possibilities how to calculate this sum. These methods are interesting and maybe more suitable for students.

The first one (Euler method) does not present strict proof. It concerns relations between roots and coefficients of finite polynomials and generalizes these relations for "infinite" polynomials.

The second one (Bernoulli method) uses product of series (not absolutely convergent).

1. Euler Method

Leonhard Euler generalized properties of polymonials to infinite series.

At first we consider a quadratic equation

$$0 = a_2x^2 + a_1x + a_0 = a_2 \cdot (x - x_1) \cdot (x - x_2)$$

and we find out relations between its coefficients a_1, a_2 and its solutions x_1, x_2

$$\frac{a_0}{a_2} = x_1 \cdot x_2, \quad -\frac{a_1}{a_2} = x_1 + x_2 \ , \quad \text{hence} \quad -\frac{a_1}{a_0} = \frac{1}{x_1} + \frac{1}{x_2} \ .$$

*Toruňská 5, 18100 Praha 8, Czech Rep.

The same is true for a cubic equation

$$0 = a_3 \cdot x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0 = a_3 \cdot (x - x_1) \cdot (x - x_2) \cdot (x - x_3) \ .$$

We have again

$$-\frac{a_0}{a_3} = x_1 \cdot x_2 \cdot x_3, \quad \frac{a_1}{a_3} = x_1 \cdot x_2 + x_1 \cdot x_3 + x_2 \cdot x_3, \quad \text{and} \quad -\frac{a_1}{a_0} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} \ .$$

The similar relation holds for equations of fourth order and so on. In general

$$-\frac{a_1}{a_0} = \frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n} \tag{1}$$

for

$$0 = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0 = a_n \cdot (x - x_1) \cdot (x - x_2) \cdot \cdots \cdot (x - x_n) \ .$$

Now we apply analogically this equality to an equation of "infinite" order. Using expansion of a function sin into power series

$$\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \dots \ , \text{ or}$$
$$\frac{\sin t}{t} = 1 - \frac{t^2}{3!} + \frac{t^4}{5!} - \frac{t^6}{7!} + \dots$$

and substitution $t := \sqrt{x}$ we can rewrite equation $\frac{\sin \sqrt{x}}{\sqrt{x}} = 0$ into

$$0 = \frac{\sin \sqrt{x}}{\sqrt{x}} = 1 - \frac{x}{6} + \frac{x^2}{120} - \frac{x^3}{1440} + \dots \ .$$

Let us remind that $a_0 = 1$ and $a_1 = -\frac{1}{6}$. Solutions of such equation are $x_k = k^2 \cdot \pi^2$, where $k \in \mathbb{N}$, as $x > 0$ and $\sqrt{x} = k \cdot \pi$.

Hence we can analogically conclude according to (1) for "infinite" polynomials

$$\frac{1}{6} = -\frac{a_1}{a_0} = \frac{1}{x_1} + \frac{1}{x_2} + \cdots = \frac{1}{\pi^2} + \frac{1}{4\pi^2} + \frac{1}{9\pi^2} + \frac{1}{16\pi^2} + \dots \quad \text{and}$$

$$\frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots \quad .$$

2. Bernoulli Method

Another method is based on products of sums

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \arctan 1 = \frac{\pi}{4} = s \tag{2}$$

which is convergent by Leibniz test. The equality (2) can be derived from sum of geometrical series

$$1 - x^2 + x^4 - x^6 + \cdots = \frac{1}{1 + x^2} \ \text{ by integration } \ x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots = \arctan x \ .$$

The main idea of the following method goes back to Nicolaus Bernoulli. Product of sequences $(a_0, -a_1, a_2, -a_3, \dots) = (1,\ -\frac{1}{3},\ \frac{1}{5},\ -\frac{1}{7}, \dots)$ gives us something like infinite matrix or infinite table

$$\begin{array}{cccccccc}
1 & -\frac{1}{3} & \frac{1}{5} & -\frac{1}{7} & \frac{1}{9} & -\frac{1}{11} & \frac{1}{13} & \cdots \\
-\frac{1}{3} & \frac{1}{9} & -\frac{1}{15} & \frac{1}{21} & -\frac{1}{27} & \frac{1}{33} & -\frac{1}{39} & \cdots \\
\frac{1}{5} & -\frac{1}{15} & \frac{1}{25} & -\frac{1}{35} & \frac{1}{45} & -\frac{1}{55} & \frac{1}{65} & \cdots \\
-\frac{1}{7} & \frac{1}{21} & -\frac{1}{35} & \frac{1}{49} & -\frac{1}{63} & \frac{1}{77} & -\frac{1}{91} & \cdots \\
\frac{1}{9} & -\frac{1}{27} & \frac{1}{45} & -\frac{1}{63} & \frac{1}{81} & -\frac{1}{99} & \frac{1}{117} & \cdots \\
-\frac{1}{11} & \frac{1}{33} & -\frac{1}{55} & \frac{1}{77} & -\frac{1}{99} & \frac{1}{121} & -\frac{1}{143} & \cdots \\
\frac{1}{13} & -\frac{1}{39} & \frac{1}{65} & -\frac{1}{91} & \frac{1}{117} & -\frac{1}{143} & \frac{1}{169} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}.$$

We shall denote by A_n, $n = 0, 1, 2, \dots$ the sum of diagonals and B_n, $n = 1, 2, \dots$ sum of skew diagonals according to the folloving schemas

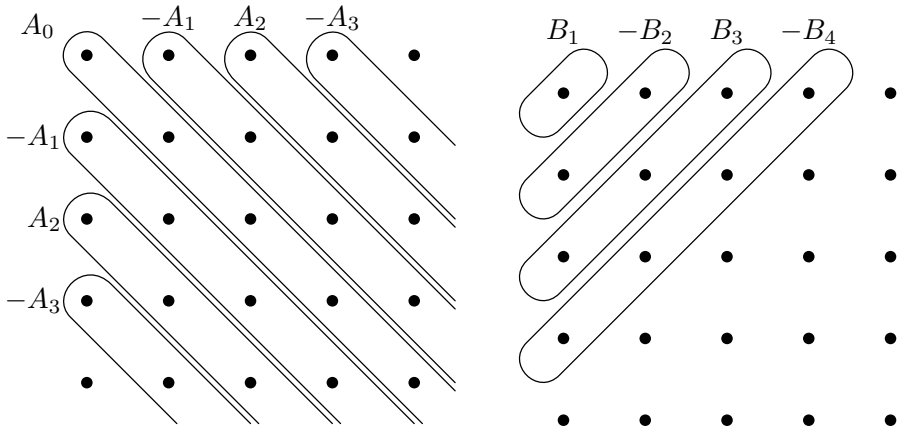


Fig. 1. The sums of diagonals and skew diagonals.

We want to investigate the sum A_0. It holds $2 \cdot A_n = B_n$ for $n = 1, 2, \dots$. We can add that infinite matrix by squares (partial sums Q_n), then we can add the diagonals (partial sums S_n) and the skew diagonals (partial sums T_n).

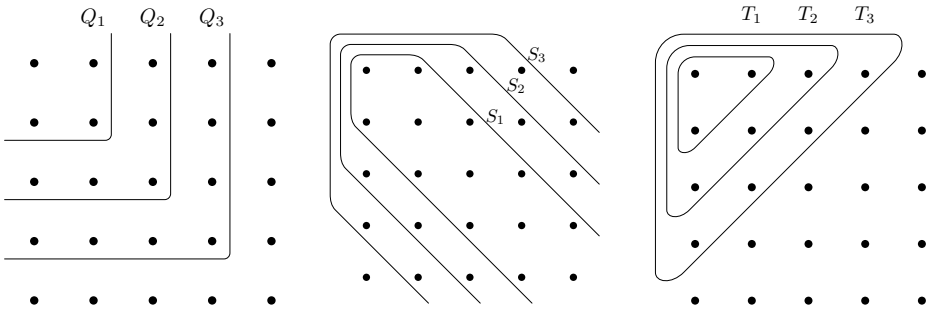


Fig. 2. The possibilities of summation.

All three methods of summation give us the same number, which equals to $s^2 = (\frac{\pi}{4})^2$, hence from

$$S_n = A_0 - 2A_1 + 2A_2 - 2A_3 + \cdot + (-1)^n \ldots 2A_n = \\ = A_0 - B_1 + B_2 - B_3 + \cdot + (-1)^n \ldots B_n = A_0 - T_n$$

we have

$$s^2 = A_0 - s^2 \quad \text{and so} \quad A_0 = 2s^2 = \frac{\pi^2}{8} \ .$$

Now more precisely and correctly. The series (2) is not absolutely convergent. That is why we cannot use Mertens' theorem and we must use other ways for proof.

It is clear that the series

$$A_0 = a_0^2 + a_1^2 + a_2^2 + a_3^2 + \cdots = 1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \ldots \ ,$$

is convergent by the integral test or Cauchy condensation test for instance. The partial sums by squares are convergent and

$$Q_n = (1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1})^2 \to s^2 = \frac{\pi^2}{16} \ .$$

2.1. *Relations between the sums A_n and B_n and convergence of them*

At first we shall consider sums according to the diagonals A_n.

We can use the decomposition of fraction

$$\tfrac{2}{3} = 1 - \tfrac{1}{3} \, , \ \tfrac{2}{15} = \tfrac{1}{3} - \tfrac{1}{5} \, , \ \tfrac{4}{5} = 1 - \tfrac{1}{5} \, , \ \tfrac{4}{21} = \tfrac{1}{3} - \tfrac{1}{7} \, , \ldots$$

for the series A_1, A_2 etc.

$$A_1 = \tfrac{1}{3} + \tfrac{1}{15} + \tfrac{1}{35} + \cdots = \tfrac{1}{2} \cdot (\underbrace{1 - \tfrac{1}{3}}_{=\frac{2}{3}} + \underbrace{\tfrac{1}{3} - \tfrac{1}{5}}_{=\frac{2}{15}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{7}}_{=\frac{2}{35}} + \dots) = \tfrac{1}{2} \cdot 1 \ ,$$

$$\begin{aligned} A_2 &= \tfrac{1}{5} + \tfrac{1}{21} + \tfrac{1}{45} + \tfrac{1}{77} + \cdots = \\ &= \tfrac{1}{4} \cdot (\underbrace{1 - \tfrac{1}{5}}_{=\frac{4}{5}} + \underbrace{\tfrac{1}{3} - \tfrac{1}{7}}_{=\frac{4}{21}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{9}}_{=\frac{4}{45}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{11}}_{=\frac{4}{77}} + \dots) = \tfrac{1}{4} \cdot (1 + \tfrac{1}{3}) \ , \end{aligned}$$

$$\begin{aligned} A_3 &= \tfrac{1}{7} + \tfrac{1}{27} + \tfrac{1}{55} + \tfrac{1}{91} + \cdots = \\ &= \tfrac{1}{6} \cdot (\underbrace{1 - \tfrac{1}{7}}_{=\frac{6}{7}} + \underbrace{\tfrac{1}{3} - \tfrac{1}{9}}_{=\frac{6}{27}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{11}}_{=\frac{6}{55}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{13}}_{=\frac{4}{91}} + \dots) = \tfrac{1}{6} \cdot (1 + \tfrac{1}{3} + \tfrac{1}{5}) \end{aligned}$$

and so on. In general we have

$$A_n = \frac{1}{2n} \cdot (1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}) \ .$$

There are also interesting relations between sums B_n and A_n

$$\begin{aligned} 2A_1 &= 1 = B_1 \ , \\ 4A_2 &= 1 + \tfrac{1}{3} = \tfrac{2}{3} + \tfrac{2}{3} = 2 \cdot (\tfrac{1}{3} + \tfrac{1}{3}) = 2B_2 \ , \\ 6A_3 &= 1 + \tfrac{1}{3} + \tfrac{1}{5} = (1 + \tfrac{1}{5}) + \tfrac{1}{3} = 3 \cdot (\tfrac{2}{5} + \tfrac{1}{9}) = 3 \cdot (\tfrac{1}{5} + \tfrac{1}{9} + \tfrac{1}{5}) = 3B_3 \ , \\ 8A_4 &= 1 + \tfrac{1}{3} + \tfrac{1}{5} + \tfrac{1}{7} = (1 + \tfrac{1}{7}) + (\tfrac{1}{3} + \tfrac{1}{5}) = \\ &= 4 \cdot (\tfrac{2}{7} + \tfrac{2}{15}) = 4 \cdot (\tfrac{1}{15} + \tfrac{1}{7} + \tfrac{1}{7} + \tfrac{1}{15}) = 4B_4 \ , \\ 10A_5 &= 1 + \tfrac{1}{3} + \tfrac{1}{5} + \tfrac{1}{7} + \tfrac{1}{9} = (1 + \tfrac{1}{9}) + (\tfrac{1}{3} + \tfrac{1}{7}) + \tfrac{1}{5} = \\ &= 5 \cdot (\tfrac{2}{9} + \tfrac{2}{21} + \tfrac{1}{25}) = 5 \cdot (\tfrac{1}{9} + \tfrac{1}{21} + \tfrac{1}{25} + \tfrac{1}{21} + \tfrac{1}{9}) = 5B_5 \ , \end{aligned}$$

and so on, in general

$$2A_n = B_n \ .$$

The sequence A_n is decreasing and $\lim A_n = 0$ because of well known inequality

$$\limsup \frac{a_1 + a_2 + \cdots + a_n}{n} \leq \limsup a_n \ .$$

So the series $\sum (-1)^{n+1} A_n$ is convergent by Leibniz test and we can denote

$$\sum_{n=1}^{\infty} (-1)^{n+1} A_n = A \quad (\text{or} \quad \lim S_n = A \ ) \ .$$

As $B_n = 2A_n$ holds, the same is true for sum $\sum (-1)^{n+1} B_n$ and

$$\sum_{n=1}^{\infty} (-1)^{n+1} B_n = 2A \quad (\text{or} \quad \lim T_n = 2A) \ .$$

2.2. *Sums Q_n and S_n*

Now we shall consider the sums Q_n and S_n. It holds

$$S_n = A_0 - 2A_1 + 2A_2 - 2A_3 + \cdots + 2\cdot(-1)^n A_n \to A_0 - 2A\ .$$

Now we shall focus our interest on the difference of S_n and T_n and we define sequences A_{nk} and C_n

$$-(S_1 - Q_1) = 2\underbrace{(A_1 - \tfrac{1}{3})}_{=A_{11}} - \underbrace{(\tfrac{1}{25} + \tfrac{1}{49} + \dots)}_{=C_2} \quad \text{where}$$

$$A_{11} = \tfrac{1}{15} + \tfrac{1}{35} + \tfrac{1}{63} + \cdots = \tfrac{1}{2}(\underbrace{\tfrac{1}{3} - \tfrac{1}{5}}_{=\frac{2}{15}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{7}}_{=\frac{2}{35}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{9}}_{=\frac{2}{63}} + \dots) = \tfrac{1}{2}\cdot\tfrac{1}{3}\ ,$$

$$S_2 - Q_2 \ = 2\underbrace{(A_2 - \tfrac{1}{5})}_{=A_{22}} - \underbrace{(A_1 - \tfrac{1}{3} - \tfrac{1}{15})}_{=A_{21}} + \underbrace{\tfrac{1}{49} + \tfrac{1}{81} + \dots}_{=C_3} \quad \text{where}$$

$$A_{22} = \tfrac{1}{21} + \tfrac{1}{45} + \tfrac{1}{77} + \cdots = \tfrac{1}{4}(\underbrace{\tfrac{1}{3} - \tfrac{1}{7}}_{=\frac{4}{21}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{9}}_{=\frac{4}{45}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{11}}_{=\frac{4}{77}} + \dots) = \tfrac{1}{4}\cdot(\tfrac{1}{3} + \tfrac{1}{5})\ ,$$

$$A_{21} = \tfrac{1}{35} + \tfrac{1}{63} + \tfrac{1}{99} + \cdots = \tfrac{1}{2}(\underbrace{\tfrac{1}{5} - \tfrac{1}{7}}_{=\frac{2}{35}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{9}}_{\frac{2}{63}} + \underbrace{\tfrac{1}{9} - \tfrac{1}{11}}_{=\frac{2}{99}} + \dots) = \tfrac{1}{2}\cdot\tfrac{1}{5}\ ,$$

$$-(S_3 - Q_3) = 2\underbrace{(A_3 - \tfrac{1}{7})}_{=A_{33}} - 2\underbrace{(A_2 - \tfrac{1}{5} - \tfrac{1}{21})}_{=A_{32}} + 2\underbrace{(A_1 - \tfrac{1}{3} - \tfrac{1}{15} - \tfrac{1}{35})}_{=A_{33}} -$$
$$- \underbrace{(\tfrac{1}{81} + \tfrac{1}{121} + \dots)}_{=C_4} \quad \text{where}$$

$$A_{33} = \tfrac{1}{27} + \tfrac{1}{55} + \tfrac{1}{91} + \cdots = \tfrac{1}{6}(\underbrace{\tfrac{1}{3} - \tfrac{1}{9}}_{=\frac{6}{27}} + \underbrace{\tfrac{1}{5} - \tfrac{1}{11}}_{=\frac{6}{57}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{13}}_{=\frac{6}{91}} + \dots) =$$
$$= \tfrac{1}{6}\cdot(\tfrac{1}{3} + \tfrac{1}{5} + \tfrac{1}{7})\ ,$$

$$A_{32} = \tfrac{1}{45} + \tfrac{1}{77} + \tfrac{1}{117} + \cdots = \tfrac{1}{4}(\underbrace{\tfrac{1}{5} - \tfrac{1}{9}}_{=\frac{4}{45}} + \underbrace{\tfrac{1}{7} - \tfrac{1}{11}}_{=\frac{4}{77}} + \underbrace{\tfrac{1}{9} - \tfrac{1}{13}}_{=\frac{4}{117}} + \dots) = \tfrac{1}{4}\cdot(\tfrac{1}{5} + \tfrac{1}{7})\ ,$$

$$A_{31} = \tfrac{1}{63} + \tfrac{1}{99} + \tfrac{1}{143} + \cdots = \tfrac{1}{2}(\underbrace{\tfrac{1}{7} - \tfrac{1}{9}}_{=\frac{2}{63}} + \underbrace{\tfrac{1}{9} - \tfrac{1}{11}}_{=\frac{2}{99}} + \underbrace{\tfrac{1}{11} - \tfrac{1}{13}}_{=\frac{2}{143}} + \dots) = \tfrac{1}{2}\cdot\tfrac{1}{7}\ ,$$

and so on. In general

$$(-1)^n\cdot(S_n - Q_n) = 2\cdot A_{nn} - 2\cdot A_{n\,n-1} + \cdots + (-1)^{n-1}\cdot 2\cdot A_{n1} + (-1)^n\cdot C_{n+1}$$
$$\text{where } A_{nm} = \frac{1}{2m}\Big(\frac{1}{2n+1} + \frac{1}{2n-1} + \cdots + \frac{1}{2(n-m)+3}\Big)\ ,$$
$$C_n = \frac{1}{(2n+1)^2} + \frac{1}{(2n+3)^2} + \dots \quad \text{for} \quad 1 \le m \le n\ .$$

Simple calculation give us the folloving inequalities and limits

$$A_{n1} \le A_{n2} \le \cdots \le A_{n\,n-1} \le A_{nn} \quad \text{and}$$
$$0 \le A_{nn} - A_{n\,n-1} + \cdots + (-1)^{n-1} \cdot A_{n1} \le A_{nn} \le A_n \to 0$$

$$C_n \to 0 \ .$$

From

$$|S_n - Q_n| \le 2 \cdot \underbrace{|A_{nn} - A_{n\,n-1} + \cdots + (-1)^{n-1} \cdot A_{n1}|}_{\le |A_{nn}| \le A_n \to 0} + \underbrace{|C_{n+1}|}_{\to 0}$$

we have

$$S_n - Q_n \to 0$$

and hence the limit of both partial sums must approach the same value and

$$A_0 - 2A = s^2 \ . \tag{3}$$

2.3. *Sums T_n and Q_n*

Inequalities between sums T_n and Q_n

$$T_1 = B_1 - B_2 \le Q_1 \le B_1 - B_2 + B_3 = T_2 \ ,$$
$$T_3 = B_1 - B_2 + B_3 - B_4 \le Q_3 \le$$
$$\le B_1 - B_2 + B_3 - B_4 + B_5 - B_6 + B_7 = T_6 \ ,$$
$$T_5 = B_1 - B_2 + B_3 - B_4 + B_5 - B_6 \le Q_5 \le$$
$$\le B_1 - B_2 + B_3 - B_4 + B_5 - B_6 + B_7 - B_8 + B_9 - B_{10} + B_{11} = T_{10} \ ,$$

generally

$$T_n = B_1 - B_2 + \cdots + (-1)^n \cdot B_{n+1} \le Q_n \le B_1 - B_2 + \cdots + B_{2n+1} = T_{2n}$$

for any odd n give us

$$\lim T_n \le \lim Q_n \le \lim T_n, \quad 2A \le s^2 \le 2A \quad \text{and} \quad 2A = s^2 \ . \tag{4}$$

2.4. *Conclusion*

As we remind (3) and (4) we have

$$s^2 = A_0 - 2A = A_0 - s^2 \ , \quad \text{hence} \quad A_0 = 2s^2 = \frac{\pi^2}{8}$$

and

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = A_0 = \frac{\pi^2}{8} \ .$$

The investigated series can be divided into odd and even terms

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} + \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{\pi^2}{8} + \frac{1}{4} \sum_{n=1}^{\infty} \frac{1}{n^2}$$

and we can conclude

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \ .$$

References

Konrad Knop (1957). *Theory and Application of Infinitive Series*, London.
Peter Ekkehard Kopp (2005). Analysis. *Bulletin of Applied Mechanics*, **3**: 171-179.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3 No. 1, pp. 16 - 23, 2010

BOUNDED INVERSE RELATIONAL FUNCTION

Chandan Kumar
Department of Civil Engineering,
Delhi College of Engineering,
Bawana Road, Delhi-110042, India,
chandankumardce@gmail.com/yahoo.co.in

In the present paper,we find the relationship between three functions, Bijective function, Inverse of the function and Identity function and their behaviour in the bounded region.

Keywords: Mean Value Theorem, Bounded Inverse Theorem, continuity, Differentiability, Monotonic function, Graphs.

1. Introduction

An Identity function is a function that always returns the same value that was used as its argument. In terms of equation, the function is given by $f(x) = x$. A Bijective function is a function f from set X to set Y with the property that for every y in Y, there is exactly one x in X, such that $f(x) = y$ and no unmapped element remains in both X and Y. Bijective function plays a fundamental role in many areas of mathematics, for instance in the definition of Isomorphism, Permutation group, Projective map and many others. Let f be a function whose domain is the set X, and whose co-domain is the set Y. Then if it exists, the inverse of f is the function f^{-1} with domain Y and co-domain X defined by: $f(x) = y$ and $f^{-1}(y) = x$ i.e a function is invertible if and only if its inverse relation is a function in which case the inverse relation is the inverse function. The inverse relation is the relation obtained by switching x and y every where.In this paper we study the relationship between Bijective function, Inverse of the function and Identity function, in a interval. Generally theorems are derived on individual function, while in this paper we derive results on the combination of functions.

2. Theorem.

Let there be any function say f : [a,b] $\rightarrow$ [d,e],
Then if :
(1) The function f(x) is continuous in [a,b]
(2) The function f(x) is differentiable in (a,b)

(3) The function f(x) is Bijective in [a,b]
(4) The sign of both a , b and f(a) , f(b) are opposite
(5) Either : $|f(a)| < |a|$ and $|f(b) < |b|$ or $|f(a)| > |a|$ and $|f(b)| > |b|$

Let $f^{-1}(x) = g(x)$
Then there exists at least one point c such that :

$$f(c) = g(c) = c \quad and \tag{1}$$

$$f'(c) \times g'(c) = 1 \;\; for \;\; f'(c) = g'(c) \neq 0 \tag{2}$$

where c is the point of intersection of function $f(x), g(x)$ and $y = x$,
and g : [d,e] $\rightarrow$ [a,b]

$$For \qquad f(c) = g(c) = c \tag{3}$$

2.1. *Proof.1:*

2.1.1. *Case : (1)*

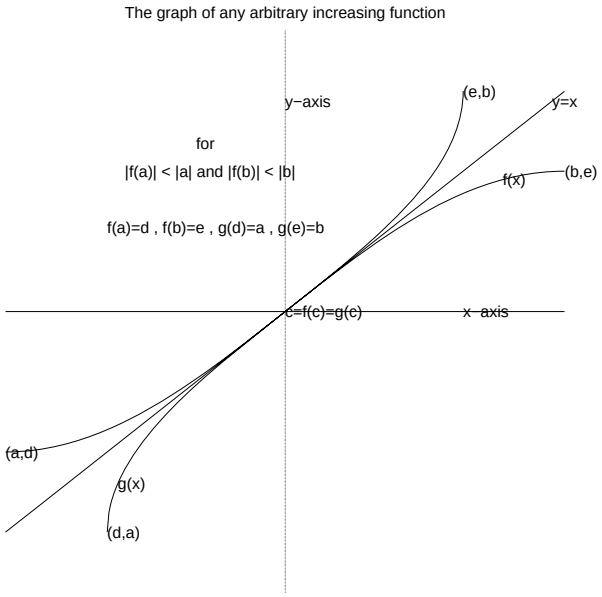
Let the function f(x) is strictly increasing for $a < b$ and $f(a) < f(b)$.Since a and b have opposite signs and f(a) and f(b) also have opposite signs. Then the co-ordinate [a,f(a)] and [b,f(b)] will lie in the opposite quadrant. When $|f(a)| < |a|$ and $|f(b)| < |b|$ then the point (a,f(a)) will lie above line y=x and the point (b,f(b)) will below line y=x. Since the co-ordinates (a,f(a)) and (b,f(b)) lie on the opposite side of the line y=x. The function f(x) is continuous between point (a,f(a)) and (b,f(b)). So there must exist atleast a single point say (k,f(k)) at the intersection of curve f(x) and line y=x. Hence k=f(k).

2.1.2. *Case : (2)*

Let the function f(x) is strictly increasing then function g(x) will also be strictly increasing. since $g(d) < g(e)$ for $d < e$ then the co-ordinates [d,g(d)] and [e,g(e)] will lie in the opposite quadrant. when $|f(a)| < |a|$ and $|f(b)| < |b|$ then $|g(d)| > |d|$ and $|g(e)| > |e|$ and the points (d,g(d)) will lie below line y=x and point (e,g(e)) will above line y=x. Since the co-ordinate (d,g(d)) and (e,g(e)) lie on the opposite side of the line y=x. The function g(x) is continuous between point (d,g(d)) and (e,g(e)). So there must exist at least a single point say (l,g(l)) at the intersection of curve g(x) and line y=x. Hence l=g(l).

2.1.3. *Case : (3)*

Let the function f(x) and its inverse g(x) are the reflection image about the line y=x and the co-ordinates of x and y are interchange by the property of inverse of the function. Let (m,f(m)) be any arbitrary point nearest to the line y=x on the curve of the function f(x) then its corresponding nearest point on the curve of the function g(x) will be (f(m),m). If the curve of the function f(x) intersect the line y=x at any



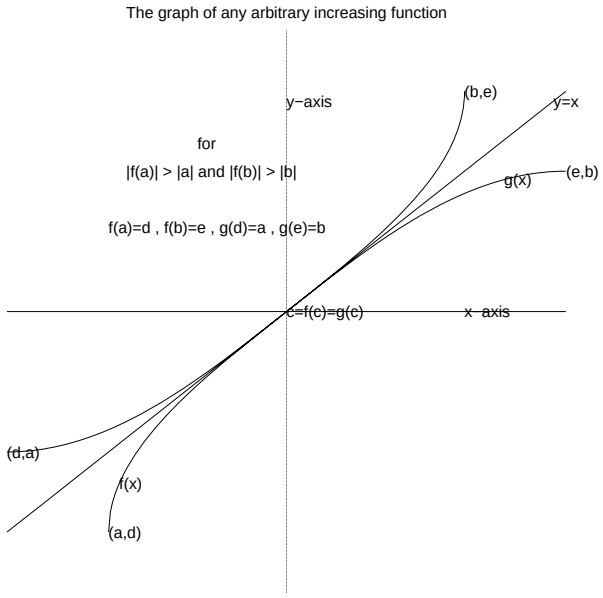
point then the inverse g(x) of the function will also intersect at the same point. So m=f(m)=g(m). From cases (1),(2) and(3) we get k=f(k)=g(k), l=f(l)=g(l) and m=f(m)=g(m). Hence f(c)=g(c)=c.

2.1.4. *Case : (4)*

Let function f(x) is strictly increasing for $a < b$, $f(a) < f(b)$. Let a and b have opposite sign and f(a) and f(b) also have opposite sign. Then the co-ordinate [a,f(a)] and [b,f(b)] will lie in the opposite quadrant. When $|f(a)| > |a|$ and $|f(b)| > |b|$ then the point (a,f(a)) will lie below line y=x and the point (b,f(b)) be will above line y=x. The co-ordinate (a,f(a)) and (b,f(b)) lie on the opposite side of the line y=x. Since the function f(x) is continuous between point (a,f(a)) and (b,f(b)). So there must exist at least a single point say (n,f(n)) at the intersection of curve f(x) and line y=x. Hence n=f(n).

2.1.5. *Case : (5)*

Let the function f(x) is strictly increasing then function g(x) will also be strictly increasing and $g(d) < g(e)$ for $d < e$. Let d and e have opposite sign and g(d) and g(e) also have opposite sign. Then the co-ordinate [d,g(d)] and [e,g(e)] will lie in the opposite quadrant. When $|f(a)| > |a|$ and $|f(b)| > |b|$ then $|g(d)| < |d|$ and $|g(e)| < |e|$. Then the point (d,g(d)) will lie above line y=x and the point (e,g(e)) will be below the line y=x. The co-ordinate (d,g(d)) and (e,g(e)) lie on the opposite side of the line y=x and function g(x) is continuous between point



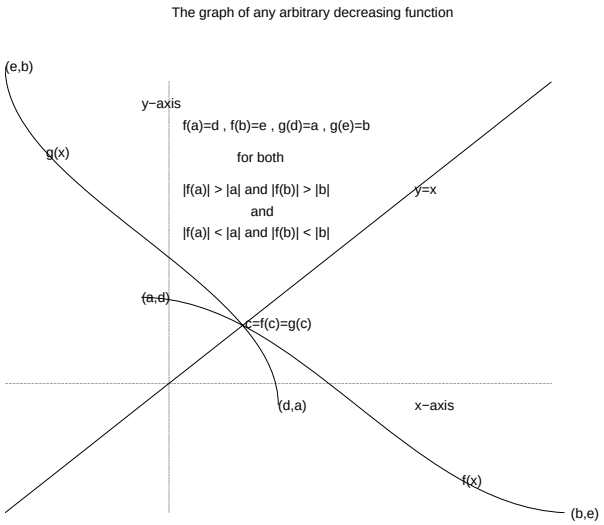
(d,g(d)) and (e,g(e)). Then there must exist at least a single point say (o,g(o)) at the intersection of curve g(x) and line y=x. So, o=g(o). From cases (3),(4) and (5) we get n=f(n)=g(n), o=f(o)=g(o) and m=f(m)=g(m). Hence f(c)=g(c)=c.

2.1.6. *Case : (6)*

Let function f(x) is strictly decreasing for $a < b$ and $f(a) > f(b)$. Let a and b have opposite sign and f(a) and f(b) also have opposite sign. Then the co-ordinate [a,f(a)] and [b,f(b)] will lie in the opposite quadrant. If $a < b$ and $f(a) > f(b)$ then a and b have opposite sign and a and f(a) will also have opposite sign. since $a < 0$ then $f(a) > 0$ and $b > 0$ then $f(b) < 0$. Since a is negative and f(a) is positive so co-ordinate [a,f(a)] will always lie above the line y=x also b is positive and f(b) is negative. Hence [b,f(b)] will always lie below the line y=x. Since the co-ordinates (a,f(a)) and (b,f(b)) lie on the opposite side of the line y=x. The function f(x) is continuous between the points (a,f(a)) and (b,f(b)). So there must exist at least one single point say (p,f(p)) at the intersection of curve f(x) and line y=x. Hence p=f(p).

2.1.7. *Case : (7)*

Let function f(x) is strictly decreasing then its inverse g(x) is also be strictly decreasing $g(e) > g(d)$ for $e < d$. If d and e have opposite sign and g(d) and g(e) also have opposite sign. Then the co-ordinates [d,g(d)] and [e,g(e)] will lie in the opposite quadrant. Since both $e < d$ and $g(e) > g(d)$ and a and b have opposite



sign for $e < 0$ then $g(e) > 0$ and $d > 0$ then $g(d) < 0$. Since e is negative and g(e) is positive so co-ordinate [e,g(e)] will always lie above the line y=x also d is positive and g(d) is negative hence [d,g(d)] will always lie below the line y=x. The co-ordinate (d,g(d)) and (e,g(e)) lie on the opposite side of the line y=x and the function g(x) is continuous between points (d,g(d)) and (e,f(e)). Hence there must exist at least a single point say (q,g(q)) at the intersection of curve g(x) and line y=x. So q=g(q). From cases (3),(6) and (7) we get p=f(p)=g(p), q=f(q)=g(q) and m=f(m)=g(m). Hence f(c)=g(c)=c

$$For \qquad f'(c) \times g'(c) = 1 \tag{4}$$

2.2. *Proof 2:*

Since both the functions f(x) and g(x) are continuous and differentiable and they intersect each other at any point on line y=x say [c,f(c)] or [c,g(c)] since f(c)=g(c)=c Now the equation of the tangent on the curve of function f(x) will be:

$$f'(c) = \frac{y-c}{x-c} \tag{5}$$

And the equation of the tangent on the curve of function g(x) will be:

$$g'(c) = \frac{y-c}{x-c} \tag{6}$$

On solving equation (5) and equation (6), we get:

$$y - x \times f'(c) + c[f'(c) - 1] = 0 \tag{7}$$

and

$$y - x \times g'(c) + c[g'(c) - 1] = 0 \tag{8}$$

Then the equation of bisectors of both the tangents is given by:

$$\frac{y - x \times f'(c) + c[f'(c) - 1]}{\sqrt{1 + [f'(c)]^2}} = + \frac{y - x \times g'(c) + c[g'(c) - 1]}{\sqrt{1 + [g'(c)]^2}} \tag{9}$$

and

$$\frac{y - x \times f'(c) + c[f'(c) - 1]}{\sqrt{1 + [f'(c)]^2}} = -\frac{y - x \times g'(c) + c[g'(c) - 1]}{\sqrt{1 + [g'(c)]^2}} \tag{10}$$

On differentiating equation (9) and equation (10) with respect to x , we get:

$$\frac{y' - f'(c)}{\sqrt{1 + [f'(c)]^2}} = +\frac{y' - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{11}$$

$$and \qquad \frac{y' - f'(c)}{\sqrt{1 + [f'(c)]^2}} = -\frac{y' - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{12}$$

2.2.1. *Case : (1)*

Since the graph of f(x) and g(x) are reflection image about line y=x. Hence one of the angle bisectors of both the tangents gives the equation of line y=x. Since the slope of line y=x is given by $y' = 1$. Substituting $y' = 1$ into (11) and (12) we get:

$$\frac{1 - f'(c)}{\sqrt{1 + [f'(c)]^2}} = +\frac{1 - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{13}$$

$$and \quad \frac{1 - f'(c)}{\sqrt{1 + [f'(c)]^2}} = -\frac{1 - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{14}$$

On squaring both sides of (13) and (14) we get:

$$\frac{[1 - f'(c)]^2}{1 + [f'(c)]^2} = \frac{[1 - g'(c)]^2}{1 + [g'(c)]^2} \tag{15}$$

On solving and simplifying equation (15), we get:

$$[1 + (f'(c))^2 - 2f'(c)] \times [1 + (g'(c))^2] = [1 + (g'(c))^2 - 2g'(c)] \times [1 + (f'(c))^2] \tag{16}$$

$$1 + (f'(c))^2 - 2f'(c) + (g'(c))^2 + [f'(c) \times g'(c)]^2 - \tag{17}$$

$$2f'(c) \times (g'(c))^2 = 1 + (g'(c))^2 - 2g'(c) + \tag{18}$$

$$(f'(c))^2 + [g'(c) \times f'(c)]^2 - 2g'(c) \times (f'(c))^2 \tag{19}$$

$$f'(c) + f'(c) \times (g'(c))^2 = g'(c) + g'(c) \times (f'(c))^2 \tag{20}$$

$$f'(c) \times g'(c)[f'(c) - g'(c)] - [f'(c) - g'(c)] = 0 \tag{21}$$

$$(f'(c) \times g'(c) - 1) \times (f'(c) - g'(c)) = 0 \tag{22}$$

$$hence \;\; f'(c) \times g'(c) = 1 \;\; for \;\; f'(c) = g'(c) \neq 0 \tag{23}$$

2.2.2. *Case : (2)*

Since one of the angle bisectors of both the tangents gives the equation of line y=x. So the other angle bisector of the tangents will be perpendicular to line y=x at that point. Since product of slope of two lines perpendicular to each other is equal to -1. So the slope of second bisector is given by $y' = -1$.

Substituting $y' = -1$ into (11) and (12) we get:

$$\frac{-1 - f'(c)}{\sqrt{1 + [f'(c)]^2}} = + \frac{-1 - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{24}$$

$$and \quad \frac{-1 - f'(c)}{\sqrt{1 + [f'(c)]^2}} = - \frac{-1 - g'(c)}{\sqrt{1 + [g'(c)]^2}} \tag{25}$$

On squaring both sides (24) and (25) we get:

$$\frac{[1 + f'(c)]^2}{1 + [f'(c)]^2} = \frac{[1 + g'(c)]^2}{1 + [g'(c)]^2} \tag{26}$$

On solving and simplifying equation (26) we get:

$$[1 + (f'(c))^2 + 2f'(c)] \times [1 + (g'(c))^2] = [1 + (g'(c))^2 + 2g'(c)] \times [1 + (f'(c))^2] \tag{27}$$

$$1 + (f'(c))^2 + 2f'(c) + (g'(c))^2 + [f'(c) \times g'(c)]^2 + \tag{28}$$

$$2f'(c) \times (g'(c))^2 = 1 + (g'(c))^2 + 2g'(c) + \tag{29}$$

$$(f'(c))^2 + [g'(c) \times f'(c)]^2 + 2g'(c) \times (f'(c))^2 \tag{30}$$

$$f'(c) + f'(c) \times (g'(c))^2 = g'(c) + g'(c) \times (f'(c))^2 \tag{31}$$

$$f'(c) \times g'(c)[f'(c) - g'(c)] - [f'(c) - g'(c)] = 0 \tag{32}$$

$$(f'(c) \times g'(c) - 1) \times (f'(c) - g'(c)) = 0 \tag{33}$$

$$hence \;\; f'(c) \times g'(c) = 1 \quad for \quad f'(c) = g'(c) \neq 0 \tag{34}$$

We conclude from case(1) and (2) that $f'(c) \times g'(c) = 1$ for all cases except $f'(c) = g'(c) \neq 0$.

2.3. *Conclusion*

We have studied the relationship between Bijective function, Inverse of the function and Identity function, in a definite interval. We have derived some standard results which may be useful in all branch of Science. Three particles (heavenly bodies, vehicles, satellites) are moving in a plane with their respective path equations $y = x$, $y = f(x)$, and $y = f^{-1}(x)$. If they satisfy the above stated conditions, then they can meet at a common point and the products of their velocities will be equal to one. This theorem is first of its kind which combines three function and gives standard results.

Acknowledgement

I would like to dedicate this paper to my father Late Shri. Mangal Das who always encouraged me in the field of research. My thanks goes to Honourable Vice Chancellor Prof. P.B. Sharma, Prof. S.K. Singh and the entire faculty of Civil Engineering Department, Delhi Technological University, for their help and to Dr. C.P. Singh of Applied Mathematics Department for his insightful suggestions. I also acknowledge the clarifying comments raised by the anonymous referees.

References

Balachandran, K., Sundaram, P., and Maki, H.,(1991):On generalized continuous maps in topological spaces, Mem. Fac. Sci. Kochi Univ. Ser. A Math.,**12**: 513.
Boisad, L.E., Regan, D.O., and Royalty, W.D.,(1988):Existence and non existence for a singular boundary valve problem, Applicable Analysis **28**: 245-246.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3 No. 1, pp. 24 - 30, 2010

NUMBERS : RATIONAL, IRRATIONAL OR TRANSCENDENTAL ?

KALLOL PAUL

Department of Mathematics, Jadavpur University
Kolkata, West Bengal, India 700032
e-mail:kalloldada@gmail.com

DIPANWITA PAUL GHOSH*

Department of Mathematics, Victoria Institution (College)
Kolkata, West Bengal 700009
e-mail: dipanwitadidi@gmail.com

We talk of rational numbers, irrational numbers, algebraic numbers, transcendental numbers and briefly describe some real numbers π, Euler's number e, Euler's Constant γ, Liouville constant α. The first two numbers π and Euler's number e are used extensively in undergraduate courses, in fact π is known to all of us from schooldays. We then mention some real numbers for which it is still not known whether they are rational, irrational or transcendental?

Keywords: Rational number, transcendental number, Liouville number.

1. Introduction

The number system of mathematics begins with the counting numbers $1, 2, 3, \ldots$ which are called natural numbers and is denoted by N. If $m, n \in N$ then we cannot always solve the equation $x + m = n$ in N and so from N we arrive at the set of integers Z in which we can solve the above type of equations. Again if $a, b \in Z$ and $a \neq 0$ then we cannot always solve the equation $ax = b$ in Z. So from Z we move to the set of rational numbers Q by taking the solutions of the above equations. Every non-empty bounded above subset of Q does not have a least upper bound in Q. We move from Q to R to ensure that every non-empty bounded above subset of R has a least upper bound in R. Moving from Q to the set of real numbers R is not as simple as moving from N to Z or from Z to Q. It requires the concept of Dedekind cut or Cauchy sequence of rationals or sequence of nested intervals. Our aim is not to talk of construction of real numbers from rational numbers in which one can define a real number as an element of a complete ordered field. One can say

*corresponding author

that N is incomplete w.r.t +, Z is incomplete w.r.t ., Q is incomplete w.r.t. $<$ and R is complete w.r.t all of them. We here talk of the real numbers which are termed as rational, irrational, algebraic, transcendental, Liouville numbers in the set of real numbers and try to look at the nature of some real numbers like Euler's number e, π and Euler's Constant γ. For the sake of clarity we first give the following definitions and then raise some pertinent questions and try to answer them.

Definition 1.1. ***Rational number***
A real number x is said to be a rational number iff it can be written in the form $x = \frac{a}{b}$ *where a,b are integers and* $b \neq 0$. *One can interpret a rational number as follows:*
The ratio of two given real numbers a and b ($b \neq 0$) *is said to be a rational number or the two numbers a and b is said to be commensurable iff there exists some integer k such that when a is divided into k equal parts it turns out that the second number b can also be divided into m such equal parts for some integer m. This interpretation becomes more clear for a positive rational number.*
With this interpretation we can say that $x = \frac{2\sqrt{3}\pi}{5\sqrt{3}\pi}$ *is a rational number, here a is divided into 2 equal parts each of length* $l = \sqrt{3}\pi$ *and b is divided into 5 such equal parts.*

Definition 1.2. ***Irrational number***
A real number x is said to be an irrational number iff it is not a rational number.

Definition 1.3. ***Algebraic number***
A real number a is said to be an algebraic number if the number a satisfies an equation of the form

$$c_n x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0 = 0$$

where the coefficients $c_n, c_{n-1}, \ldots c_0$ *are all rational numbers and* $c_n \neq 0$.
One can assume that the coefficients are all integers for, if $c_n = \frac{p_n}{q_n}$, $c_{n-1} = \frac{p_{n-1}}{q_{n-1}}, \ldots c_0 = \frac{p_0}{q_0}$ *and* $d_n = p_n q_{n-1} q_{n-2} \ldots q_1 q_0$, $d_{n-1} = q_n p_{n-1} q_{n-2} \ldots q_1 q_0$, $d_0 = q_n q_{n-1} \ldots q_1 p_0$. *Then a satisfies an equation of the form*

$$d_n x^n + d_{n-1}x^{n-1} + \ldots + d_1 x + d_0 = 0$$

where the coefficients $d_n, d_{n-1}, \ldots d_0$ *are all integers and* $d_n \neq 0$.
Thus a real number a is an algebraic number iff a satisfies a polynomial equation with rational or integral coefficients.

Definition 1.4. ***Transcendental number***
A real number x is said to be a transcendental number iff it is not an algebraic number.

Once we described the rational , irrational, algebraic and transcendental numbers we now raise the following questions and try to answer them :

Q.1. Whether $\sqrt{p}$ is rational for a prime p and more generally whether $p^{\frac{1}{n}}$ is rational for $n > 1$?

Q.2. What is the relation between rational numbers and algebraic numbers?

Q.3. Are the following numbers rational : Euler's number e, the number π, Euler's constant γ ? Are the numbers transcendental?

Q.4. What is the nature of the number $10^{-1!} + 10^{-2!} + 10^{-3!} + \ldots$? Is it irrational? Is it transcendental?

Q.5. What about the nature of the numbers of the form

$$2^{\sqrt{2}},\ 3^{\sqrt{5}},\ e^e, \alpha^\beta$$

where α and β are not algebraic ?

The answer to the first question follow from the theorem given below:

Theorem 1.5. *If the polynomial equation $d_n x^n + d_{n-1}x^{n-1} + \ldots + d_1 x + d_0 = 0$ with integral coefficients has a rational root $\frac{a}{b}$ then $a|d_0$ and $b|d_n$.*

Proof. Suppose the polynomial equation $d_n x^n + d_{n-1}x^{n-1} + \ldots + d_1 x + d_0 = 0$ with integral coefficients has a rational root $\frac{a}{b}$. Then we get

$$\begin{aligned} d_n\frac{a^n}{b^n} + d_{n-1}\frac{a^{n-1}}{b^{n-1}} + \ldots\ldots + d_1\frac{a}{b} + d_0 = 0 \\ \Rightarrow d_n a^n + bd_{n-1}a^{n-1} + \ldots + b^{n-1}d_1 a + b^n d_0 = 0 \\ \Rightarrow d_n a^n + bd_{n-1}a^{n-1} + \ldots\ldots + b^{n-1}d_1 a = -b^n d_0 \\ \Rightarrow b^n d_0 + b^{n-1}d_1 a + \ldots\ldots + bd_{n-1}a^{n-1} = -d_n a^n \end{aligned}$$

From the last two equations we get for some integers k_1 and k_2

$$ak_1 = -b^n d_0 \quad and \quad bk_2 = -d_n a^n.$$

As a and b has no common factor this shows that $a|d_0$ and $b|d_n$.

Corollary 1.6. *If the polynomial equation $x^n + d_{n-1}x^{n-1} + \ldots + d_1 x + d_0 = 0$ with integral coefficients has a rational root then it must be an integer which divides d_0.*

Proof. Let the rational root be $\frac{a}{b}$. Then $a|d_0$ and $b|1$ so that $b = 1$. Thus the rational root is an integer which divides d_0.

As an application of this corollary we see that the equation $x^n - p = 0$ has no rational root for $n > 1$ and p a prime. So $p^{\frac{1}{n}}$ is not a rational number. Further argument wil ensure that for $a \in N$ $a^{\frac{1}{n}}$ is a rational number if a is of the form b^n

for some $b \in N$. Thus $\sqrt{2}, \sqrt{5}, 7^{\frac{2}{3}}$ are all irrational numbers.

From the definition of an algebraic number it follows that every rational number $\frac{a}{b}$ is an algebraic number as the number $\frac{a}{b}$ satisfies the equation $bx - a = 0$. The converse is however not true as for a given prime p, $p^{\frac{1}{n}}$ is an algebraic number, for it satisfies the equation $x^n - p = 0$ but it is not a rational number. **The observations made here answer Q. No. 2.**

We next try to answer Q.No.3

Euler's Number e : The number e was first studied by the Swiss mathematician Euler in 1720s, although its existence was more or less implied in the work of John Napier, the inventor of logarithms in 1614. Euler was the first to use the letter e for it in 1727 (the fact that it is the first letter of his surname is coincidental). As a result e is called the Euler's number or Napier's constant. e is defined as the limit of the convergent sequence $(1 + \frac{1}{n})^n$ or as the infinite sum $\frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \ldots$. Euler gave an approximation of e upto 18 decimal places as e = 2.718281828459045235. Now its value has been calculated to 869,894,101 decimal places. We give a proof that e is irrational.
If possible let $e = \frac{h}{k}$ where h and k are integers with $k > 0$. Then clearly $k!.(e - \frac{1}{0!} - \frac{1}{1!} - \frac{1}{2!} - \ldots - \frac{1}{k!})$ is an integer but

$$\begin{aligned} 0 &< k!(e - \frac{1}{0!} - \frac{1}{1!} - \frac{1}{2!} - \ldots - \frac{1}{k!}) \\ &= \frac{1}{k+1} + \frac{1}{(k+1)(k+2)} + \frac{1}{(k+1)(k+2)(k+3)} + \ldots \\ &< \frac{1}{k+1} + \frac{1}{(k+1)^2} + \frac{1}{(k+1)^3} + \ldots \\ &= \frac{1}{(k+1)} . \frac{1}{1 - \frac{1}{k+1}} = \frac{1}{k} \\ &\leq 1 \end{aligned}$$

Thus $0 < k!(e - \frac{1}{0!} - \frac{1}{1!} - \frac{1}{2!} - \ldots - \frac{1}{k!}) < 1$. This is a contradiction and so e must be irrational □
The transcendence of e was first proved by Hermite[Hermite (1873)]. A much easier proof of transcendence of e is given in the book by Niven [Niven (2005)] which is based on a paper by Hurwitz [Hurwitz (1893)]. Anyone interested to know further about e can look at the book " e: the story of a number " by Eli Maor [Maor (2009)].

The number π : π is defined as the ratio of a circle to its diameter. Archimedes (287 B.C. - 212 B.C.) was the first to give a method of calculating π to any desired degree of accuracy by the method of exhaustion. The irrationality of π was first proved by J.H. Lambert [Lambert (1761)]. We give the proof by Niven [Niven (1947)]
.

If possible let π be rational. Then π^2 is also rational. Let $\pi^2 = \frac{a}{b}$ where a and b are positive integers. Consider

$$f : [0,1] \longrightarrow R \;\; defined \;\; by \;\; f(x) = \frac{x^n(1-x)^n}{n!} \;\; where \;\; n \in N.$$

Clearly $0 < f(x) < \frac{1}{n!}$ for all x in (0,1). Also k-th derivative $f^{(k)}(0)$ and $f^{(k)}(1)$ are all integers. Let us define another function $F : [0,1] \longrightarrow R$ by

$$F(x) = b^n \sum_{k=0}^{n} (-1)^k \pi^{2n-2k} f^{(2k)}(x).$$

Then clearly F(0) and F(1) are integers. Also

$$\frac{d}{dx}\{F'(x)\sin\pi x - \pi F(x)\cos\pi x\} = F^{(2)}(x)\sin\pi x + \pi^2 F(x)\sin\pi x = \pi^2 a^n f(x)\sin\pi x$$

$$\Rightarrow \int_0^1 \pi^2 a^n f(x)\sin\pi x dx = [F'(x)\sin\pi x - \pi F(x)\cos\pi x]_{x=0}^1 = \pi[F(1)+F(0)]$$

$$\Rightarrow \int_0^1 \pi a^n f(x)\sin\pi x dx = F(1)+F(0) - \;\; an \;\; integer$$

Now $0 < f(x) < \frac{1}{n!} \; \forall n \in N \; and \; for \; 0 < x < 1$ and so $0 < \pi a^n \int_0^n f(x)\sin\pi x dx < \frac{\pi a^n}{n!} \; \forall n \in N$.
By choosing n sufficiently large we can make $\frac{\pi a^n}{n!} < 1$ and so we get a contradiction. Thus π^2 must be irrational and so also is π □
The number π is not algebraic. Its transcedency was first proved by Lindemann [Lindemann (1882)] . A nice history of π can be found in the book " A History of π (pi) " by Petr Beckmann [Beckman (1976)] and in the book by Blatner [Blatner (1999)].

Euler's Constant γ : Euler's constant γ is defined as

$$\gamma = \lim_{n\to\infty}(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - logn).$$

The existence of the limit follows from the **Integral test**:
Let f be a positive decreasing function on $[1,\infty)$ *such that* $\lim_{x\to\infty} f(x) = 0$. *For* $n = 1, 2, \ldots$ *define* $s_n = \sum_{k=1}^n f(k), t_n = \int_1^n f(x)dx, d_n = s_n - t_n$. *Then we have*
(i) $0 < f(n+1) \le d_{n+1} \le d_n \le f(1), n = 1, 2, \ldots$
(ii) $\lim_{n\to\infty} d_n$ *exists*
(iii) $\sum_{n=1}^{\infty} f(n)$ *converges iff* $\{t_n\}$ *converges*
(iv) $0 \le d_k - \lim_{n\to\infty} d_n \le f(k)$ *for* $k = 1, 2, \ldots$
Choosing $f(x) = \frac{1}{x}$ the existence of γ is justified.
Although γ is calculated upto 1,000,000 digits it is still unknown whether γ is rational or not. γ was calculated to 16 digits by Euler in 1781. Its value is .57721566..... . If γ is a simple fraction $\frac{a}{b}$ then it is proved that $b > 10^{1242080}$. As we are not certain about the rationality of γ so we can't decide the transcendency of γ. For further information on γ one can look at the book " Gamma : Exploring Euler's

constant " by J. Havil [Havil (2009)].

Next we discuss problem 4. In 1851, the French Mathematician Liouville [Liouville (1844)]first established that transcendental numbers exist by exhibiting certain numbers which he proved to be non-algebraic. These numbers are now called Liouville numbers. A real number ξ is said to be a Liouville number if for every positive integer m there is a rational number $\frac{h_m}{k_m}$ with $k_m > 1$ such that $|\xi - \frac{h_m}{k_m}| < {k_m}^{-m}$. The number α is one such number and is called the Liouville's constant.

To answer Q.No.5 we start with Hilbert's 7th problem. In 1900 Hilbert announced a list of 23 outstanding problems. The seventh problem was to decide whether α^β is algebraic or not, given that α and β are algebraic numbers. In 1934 it was settled by Gelfond [Gel'fond (1934)] and independently by Schneider [Schneider (1934)] that α^β is transcendental (the case $\alpha = 0,\ \alpha = 1$ and $\beta =$ *rational* were excluded). So $2^{\sqrt{2}},\ 3^{\sqrt{5}}$ are transcendental . The number e^π is transcendental but it is still not known whether $e^e,\ \pi^e$ is algebraic or not. The problem like whether α^β is transcendental for α, β not algebraic are still open.
Even it is not known whether $\pi + e$, $\pi - e$, $\pi.e$, π/e are transcendental or not ?

References

Beckmann, P. (1976) . *A History of* π, St. Martin's Griffin, 3rd Ed.

Blatner,D. (1999). *The Joy of* π, Walker and Company.

Gelfond, A.O. (1934). Sur le septime problme de Hilbert, *Dokl. Akad. Nauk SSSR*, **2** : 46.

Havil, J. (2009). *Gamma: Exploring Euler's Constant*, Princeton University Press.

Hermite, C. (1873) . Sur la fonction exponentielle, *Compt. Rend. Acad. Sci. Paris*, **77** : 18-24, 74-79, 285-293.

Hurwitz, A. (1893). Beweis der Transcendenz der Zahl e, *Math. Annalen* , **43** : 220-222.

Lambert, J.H. (1761). Mmoire sur quelques proprits remarquables des quantits transcendentes circulaires et logarithmiques, *Histoire de l'Acadmie Royale des Sciences et des Belles-Lettres der Berlin* , 265-276

Lindemann, F. (1882). $\ddot{U}$ber die Zahl π , *Math Annalen*, **20**: 221-225.

Liouville, J. (1844) Sur des classes de trs tendues de quantits dont la valeur n'est ni algbrique, ni mme rductible des irrationelles algbriques, *C.R. Acad. Sci.* , **18** : 883885 ; 910911.

Maor, E. (2009). *e: The story of a number*, Princeton University Press.

Niven,I. (1947) A simple proof that π is irrational, *Bull. Amer. Math. Soc.*, **53** : 509.

Niven,I. (2005). *Irrational Numbers* , The Mathematical Association of America.

Schneider, T. (1934) Transzendenzuntersuchungen periodischer Functionen I, II, *J. Reine Angew. Math.* , **172** : 6569; 7074.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3, No. 1, pp 31 – 35, 2010

A MIXED QUADRATURE RULE FOR NUMERICAL INTEGRATION OF ANALYTIC FUNCTIONS BY USING BIRKHOFF YOUNG AND BOOLS QUADRATURE

Sanjit Ku. Mohanty
Department of Mathematics,
B.S. College, Nuahat, Jajpur
Email:sanjit_mohanty75@rediffmail.com

Rajani.B. Dash
Department of Mathematics,
Ravenshaw University, Orissa
Email:rajani_bdash@rediffmail.com

A quadrature rule has been constructed for the approximate evaluation of the integral of an analytic function over a line segment in complex plane. The rule has been numerically verified and its asymptotic error estimate has been obtained.

Key words: Quadrature rule, Asymptotic error, Analytic function, Numerical integration.

1. Introduction

There are several rules for the approximate evaluation of real definite integral.

$$\int_{-1}^{1} f(z)dz \qquad \text{.....(1.1)}$$

However there are only a few quadrature rules for evaluating an integral of type

$$I(f) = \int_{L} f(z)dz \qquad \text{.....(1.2)}$$

Where L is a directed line segment from the point Z_0-h to Z_0+h in the complex plane and $f(z)$ is analytic in certain domain Ω containing L.

Birkhoff - Young (1950) [3] have derived the following interpolatory type of quadrature rule:

$$BY(f) = (h/15)[24 f(Z_0) + 4 \{ f(Z_0 + h) + f(Z_0 - h)\} - \{ f(Z_0 + ih) + f(Z_0 - ih)\}] \quad \text{……..(1.3)}$$

and this rule is of precision five.

Lether (1976) using the transformation $Z = Z_0 + ht$, $t \in [-1,1]$ transformed the integral (1.2) to the integral

$$h \int_{-1}^{1} f\ (Z_0 + ht)\, dt\,. \quad \text{..... (1.4)}$$

and then made the approximation of this integral by applying standard quadrature rule meant for approximate evaluation of real definite integral (1.1). The rule so formed is termed as a TRANSFORMED RULE for the numerical integration of (1.2).

R.N. Das and G.Pradhan (1996) [1] have constructed a quadrature rule for the approximate evaluation of the integral (1.1) from two standard quadrature rules of different type but of equal precision, such rule is termed as MIXED QUADRATURE RULE.

In this paper we desire to construct a mixed quadrature rule of precision seven in the same vein for the approximation of the integral (1.2).

2. Formulation of the rule

For the construction of the desired rule we choose the rule (3) and the following Bools quadrature transformed rule

$$BL(f) = (h/45)[7 f(Z_0 - h) + 7 f(Z_0 + h) + 32 f(Z_0 - h/2) + 32 f(Z_0 + h/2) + 12 f(Z_0)] \quad \text{……..(2.1)}$$

Each of the rules under consideration is of precision five.

Denoting the truncation errors by E_{BY} and E_{BL} due to rules (1.3) and (2.1) respectively in approximating the integral (1.2) , we have

$$I(f) = BY(f) + E_{BY} \quad \text{.....(2.2)}$$

and

$$I(f) = BL(f) + E_{BL} \quad \text{.....(2.3)}$$

Where f is infinitely differentiable since it is assumed to be analytic in certain domain containing the line segment L. So by using Taylor's expansion the truncation errors associated with the quadrature rules under reference can be expressed as

$$E_{By}(f) = (-1/1890)h^7 f^{vi}(Z_0) - (1/226800)h^9 f^{viii}(Z_0) \quad \text{.....(2.4)}$$

and

$$E_{BL}(f) = (-1/15120)\, h^7 f^{vi}(Z_0) - (17/7257600)\, h^9 f^{viii}(Z_0) \quad \text{……(2.5)}$$

Now multiplying (2.2) and (2.3) by 1/15120 and 1/1890 respectively and subtracting the 1st one from the second we obtained

$$I(f) = [(8/7)\ BL(f) - (1/7)\ BY(f)] + [(8/7)\ E_{BL} - (1/7)\ E_{BY}]$$

$$= SM_2(f) + E_{SM2}$$

where $SM_2(f) = [(8/7)\ BL(f) - (1/7)\ BY(f)]$ .....(2.6)

is desired quadrature rule of precision seven for the approximate evaluation of $I(f)$ and the truncation error committed in this approximation is given by

$$E_{SM2} = [(8/7)\, E_{BL}(f) - (1/7)\ E_{BY}(f)] \quad \text{.....(2.7)}$$

The rule (2.6) may be called a MIXED TYPE QUADRATURE RULE as it is constructed from two different type of rules of same precision.

3. Error Analysis

Let $f(Z)$ is analytic in the disc $\Omega_R = \{ Z : |z - Z_0| \le R > |h| \}$

So that the points Z_0, $Z_0 \pm h$, $Z_0 \pm ih$ are all interior to the disc Ω_R. Now using Taylor's expansion

$$f(Z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad ; \quad a_n = \left(\frac{1}{n!}\right) f^{n}(Z_0)$$

in (2.7) we obtain after simplification

$$E_{SM2} = (h^9/226800)\ f^{viii}(Z_0) + \ldots \quad . \quad \text{.....(3.1)}$$

Thus the mixed rule SM_2 is of precision seven.

From (3.1) we have the following theorem.

Theorem 3.1 If f is assumed to be analytic in a domain Ω containing L then

$$E_{SM2}(f) = 0(h^9)$$

Error Comparison

From (2.4) and (2.5) we get

$$\left| E_{BL} \right| \leq \left| E_{BY} \right| \qquad \text{……..(3.2)}$$

Again from (3.1)

$$\left| E_{SM2} \right| \leq \left| E_{BL} \right| \qquad \text{……..(3.3)}$$

From (3.2) and (3.3) $\left| E_{SM2} \right| \leq \left| E_{BL} \right| \leq \left| E_{BY} \right|$

4. Numerical Verification

Let us approximate the value of the following integral using BY(f), BL (f) and $SM_2(f)$ quadrature rule

Consider $I_1 = \int_{-i}^{i} e^z dz$ and $I_2 = \int_{-i/2}^{i/2} \cos z dz$

Table for comparison of values of integrals I_1 and I_2

Sl.No	Quadrature Rules	Approximation(I_2)	Approximation(I_1)
1	BY (f)	i (1.6824171)	i(1.0421948)
2	BL (f)	i (1.6828781)	i(1.0421911)
3	SM_2 (f)	i (1.682944)	i(1.0421906)
4	Exact Value	i (1.6829420)	i(1.0421906)

5. Conclusion

From the table it is evident that the mixed quadrature rule $SM_2(f)$ is giving us better result than each of the BL(f) (Bool's rule) and BY(f) (Birkhoff-Young's rule).This conforms the theoretical result that $SM_2(f)$ is of higher precision (precision 7) than those of BY(f) and BL(f) (each having precision 5).

References

[1]. Das, R.N. and Pradhan, G. (1996) A mixed quadrature rule for numerical integration of analytic *Functions. Bull Cal. Math. Soc. 89, 37-42 (1997)*

[2]. Acharya B.P. and Das R.N. (1983*)Compound Birkhoff - young rule for numerical integration of analytic functions.Int. J. Math. Educ. Sci. Technol 14, 1, 101*

[3]. Birkhoff, G and young, D (1950) *Numerical quadrature of analytic and harmonic functions. J. Maths Phys. 29,217.*

[4]. Das, R.N. and Pradhan, G. (1996) *A mixed qudrature rule for approximate evaluation of real definite integrals, Int. J. Math. Educ. Sci. Technol*

[5]. Mohanty S. K and Dash R. B(2008*). A mixed quardature rule for numerical integation of analytic functions, Bulletin of Pure and Applied sciences, Newdelhi.Vol-27E(No.2)P.369-372.*

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3, No. 1, pp 36 – 42, 2010

ETHNOMATHEMATICS
(A Mathematical Game in Hausa Culture)

Mohammed Waziri Yusuf
Faculty of Science, Department of Mathematics University Putra,Malaysia
waziri@math.upm.edu.my
Ibrahim Saidu
Faculty of Computer Science and Information Technology, University Putra,Malaysia.
Harmach1974@yahoo.com
Aisha Halliru
Department of Mathematical sciences, Bayero University Kano,(B.U.K) Nigeria

Abstract
Before the coming of Western education, Hausa people in northern Nigeria used mathematics in sorting, ordering, measuring, timing and weighing in their day to day activities. Featuring prominently in this culture are the traditional games played by children or adults with the aim of deriving lessons. These games are meant to shape individual brain. The primary concern of this paper is on the games that involves calculations and to highlight therein the existence of algebra, set theory, coordinate geometry, arithmetic progression and geometry progression in Hausa culture. Furthermore the paper will correct the erroneous believe that there was no mathematics in Hausa land/northern Nigeria before the coming of the Western education and address those who believe mathematics only starts and ends in the classroom, thereby divorcing the rich cultural elements from the use of mathematics.

Keywords: Hausa culture; Mathematical game; Western education.

1. Introduction

Hausa is the largest and dominate language in northern Nigeria. The culture of Hausa people has over shadow almost all the other cultures in the region, up to the level of including it into National curriculum as a compulsory subject in primary and secondary education. Hausa culture is like any other culture, it consists of variety of games played by children or adult, with the aim of deriving lesson from them. In Hausa culture elders shapes younger individuals brains by storytelling and question and answer, they organize competition with the hope of getting the best out of the participants. The mathematical games in northern Nigeria have become hot games played by all categories of people.
The traditional believes of Hausa people in northern Nigeria about mathematics are:

i. There is no mathematics in Hausa culture.
ii. Mathematics starts and ends only in classroom.
iii. Mathematics has no relevance to Hausa people.

The three above believes are what motivate us to write this paper, so that we can wash away the assumption made by our people.

However there are quite a few researches which attempt to survey some of those games on General Hausa culture. Not knowing that, concentrating on these games can

improve the teaching and learning mathematics. In this paper we made a survey on some Hausa games that involves mathematics, and translate those verbal Hausa statements to mathematical expression and solve them. The paper shows the position of algebra, set theory, trigonometric, coordinate Geometry, Arithmetic Progression and Geometric Progression, e.t.c as tools of analysis for Hausa games. The paper also proves the existence of mathematical knowledge in Hausa land before the coming of western education. Moreover the paper shapes itself from the fact that Hausa games require intellectual, rational and analytical thinking, which is purely mathematics. Furthermore, Hausa people use mathematics in sorting, ordering, measuring and weighing in their daily activities.

According to [Habibu (2004)] Hausa game shapes individuals (children and adults) intellectual reasoning capabilities. Hausa game is a conceptualize way of developing cognitive development of individuals [Yahaya (1997)]. [Habibu (2004)] Identifies features and the types of leisure these games encourage in the Hausa community. [Habibu (2004)] Asserts that, it is a well known fact in Hausa culture, individuals are often for lacking "tsari" or "Lissafi". These two concepts are mathematics. Learning mathematics is not just about acquiring and mastering computational and problem solving techniques, or solely about understanding definitions, arguments and proofs. In addition to all of these things, it also involves you reconstructing the thinking of work of other mathematicians.[Heather (2003)] Also [Kanh (2003)] confirms that Learning mathematics requires you to develop ways of thinking mathematically. [Jonh (1998)] Asserts, there is a perception that mathematics is an effective tool for analyzing, examining and verifying truth. Human life remains incomplete without mathematics [James (1982)]. According to [Musa (2007)] mathematics is the construction of knowledge with regard to the qualitative and quantitative relationships of space and time. Mathematics is human activity that deals with patterns, problem solving, logical thinking, and so on, with the aim of understanding the world [Glorin (1980)]. Mathematics is a universal subject in whom every culture has its concept of numbers and the idea that 1 + 1 = 2, no matter how technologically advanced the culture. The universality notion of mathematics is further reinforced by the fact that it was invented all over the world, in a multitude of places' and different times, with little or no contact amongst its creator. According to [John (1998)] Plato proclaims that mathematics is a reliable tool for pursuing truth.

Ethnomathematics refers to the study of mathematical practices of specific cultural groups in the course of dealing with their environmental problems and activities [Glorin (1980)] and [Ascher (1991)]. The Prefix "ethno" refers to identifiable cultural groups, such as notional tribal societies, professional classes etc. and includes their language and daily practices. "mathema" here means to explain, understand and manage reality specifically by counting, measuring, classifying, ordering and modeling patterns arising in the environment. The suffix "ticks" means art to technique. According to [John (1998)] ethnomathematics is the study of mathematical techniques used by identifiable cultural groups in understanding, explaining, and managing problems and activities arising in their own domain. Ethnomathematics also refers to any form of cultural knowledge, or social activity characteristics of a social and/or cultural group that can be recognized by other groups [Louis (1986)]. For example, the manner professional basket ball players estimate angles and distances differs greatly from the corresponding manner used by truck drivers. Both professional basketball players and truck drivers are re-identifiable cultural groups that use mathematics in their daily work. They have their own

language and specific way of obtaining the estimate and ethno mathematicians study their techniques. Culture refers to a set of norms, beliefs, and values that are common to a group of people who belong to the same ethnicity [James (1982)].
The paper is arranged as follows: Procedure of the Game in section 2, analytical solution to the problems in section 3 and Conclusion in section 4.

2. Procedure of the Game

The game is of two forms namely:

(i) Say your opinion.
(ii) Get the actual answer.

The primary concern is on the games that involve mathematical formulation and solution. Because of the scanty literature on the topic, some games were sourced from [waziri et al (2010)] and broadcasting stations [Rabiu (2006), Tandarki (2007) , Musa (2007) and Auwalu (2003)] .

3. The Games

Game 1: You are asked to bring out two oranges from a garden. The garden has ten gates with gateman each. Whatever number of oranges you have it should be divided equally with each of the gateman. How many oranges are you to carry?

Solution

Let x_1 be the number of orange at 1^{st} gate.

Let x_2 be the number of orange at 2^{nd} gate.

Let x_3 be the number of orange at 3^{th} gate.

Let x_4 be the number of orange at 4^{th} gate.

Let x_5 be the number of orange at 5^{th} gate.

Let x_6 be the number of orange at 6^{th} gate.

Let x_7 be the number of orange at 7^{th} gate.

Let x_8 be the number of orange at 8^{th} gate.

Let x_9 be the number of orange at 9^{th} gate.

Let x_{10} be the number of orange at 10^{th} gate

Therefore we have this sequence $x_1, x_2, x_3 \ldots\ldots x_{10}$

The goal is to find the total number of oranges to carry at the initial point

Since at 1^{st} gate we are expected to come with 4, 2^{nd} gate with 8, 3^{rd} gate with 16, and so on up to 10^{th} gate.

Hence we can generate the following sequence 4,8,16,…….

The above sequence is a Geometric sequence.

Where

$a = 4$, $r = 2$,and $n = 10$

We need to look for the 10^{th} term. Using the below formula.

$$T_{10} = ar^{n-1} \quad (1)$$

Therefore the number of oranges needed at the 10^{th} gate is 2048

Analysis

4 8 16 32 64 128 256 512 1024 2048

1^{st} 2^{nd} 3^{rd} 4^{th} 5^{th} 6^{th} 7^{th} 8^{th} 9^{th} 10^{th}

Game 2: A father tells his son that, a spider has 8 legs, a cockroach has 6 legs and 4 wings and a grasshopper has 6 legs and 2 wings. 18 of these three kinds of insects are in a cage while the total number of legs and wings of these three kinds of insects are 118 and 40 respectively. What is the number of each kind in the cage?

Solution

The goal is to find the number of spider, cockroach and grasshopper in the cage.

Let x be the number of spider

Let y be the number of cockroach

Let z be the number of grasshopper

Observe that the total number of insects in the cage is 18, therefore we have

$$x + y + z = 18 \quad \text{(i)}$$

And also observe that the total number of legs and wings are also given.

Therefore we have

$$8x + 6y + 6z = 118 \quad \text{(ii)}$$

And

$$0x + 4y + 2z = 40 \quad \text{(iii)}$$

When we solve the three equations simultaneously we obtained the solution as $x = 5 \;\; y = 7 \;\; \text{and} \;\; z = 6$

Therefore the number of spider in the cage is 5, The number of cockroach in the cage is 7 and the number of grasshopper in the cage is 6.

Analysis

- 5 + 7 + 6 = 18 , we have18 insects .
- 8 (5) +6(7) + 6(6) = 118,we have 118 legs.
- 4(7) + 2(6) = 40 , we have 40 wings.

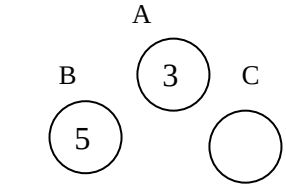
Game 3: A man has a bag container containing 8 measure of maize. A buyer come to buy half of this maize (4 measures) and he has 2 empty bag containers with the capacities 5 measure and 3 measure respectively. Without weighing, how can you put 4 measure of maize into container of 5 capacity measure?

Solution.

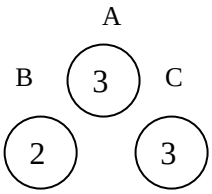
Let the container of capacity 8 measure be A and let the container of capacity 5 measure be B, let the container of capacity 3 measure be C.

And none of this container will contain more than its capacity.

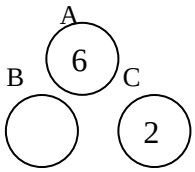
STEP 1: Put 5 measures into container B, and then we have 3 measures in container A and 5 measures in container B while C is empty



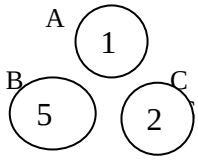
STEP 2: Put 3 measures in container C from container B now we have 3 measures in container A, 2 measures in container B and 3 measures in container C



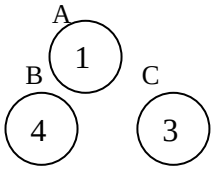
STEP 3: Return 3 measures from container C to container A and 2 measure from container B to container C now we have 6 measure in A , 2 measure in container C and container B is empty.



STEP 4: Put 5 measures to container B from container A then we have 1 measure in container A, 5 Measure in container B and 2 measure in container C



STEP 5: Since container C cannot contain more than 3 measures and it contain 2 measures, remain 1 measure, therefore from container B we fill up container C, so that container C contain 3 measures and B contain 4 measures.



Game 4: There are 8 ducks in front of 8 other ducks.
There are 8 ducks behind 8 other ducks.
There are 8 ducks beside 8 other ducks.

How many ducks are there altogether?

Solution
Consider the diagram below:

Duck 1	Duck 2	Duck 3	Duck 4
Duck 5	Duck 6	Duck 7	Duck8
Duck 9	Duck 10	Duck 11	Duck 12
Duck13	Duck14	Duck 15	Duck16

- In front of Duck 1, 2, 5,6,9,10,13 and Duck 14 there are Duck 3,4,7,8,11,12,15 and Duck 16. Therefore in front of 8 Ducks there are 8 Ducks.
- Behind Duck 3,4,7,8,11,12,15 and Duck 16 there are Duck 1, 2, 5,6,9,10,13 and Duck 14. Therefore behind 8 Ducks there are 8 Ducks.
- Beside Duck 1, 2, 3, 4,5,6,7 and Duck 8 there are Duck 9,10,11,12,13,14,15 and Duck 16. Therefore in front of 8 Ducks there are 8 Ducks.
- The total Ducks altogether are **sixteen (16).**

Game 5: A father has a certain numbers of onions and tomatoes, but he doesn't know the actual number of them. Then he asked his son to tell him the number of onions and tomatoes at hand, the boy said if you double the sum of the onions and tomatoes the number is 26, but if you double the product of onions and tomatoes will be 80. What is the actual number of onions and tomatoes that the father got?

Solution

- Let x to be the number of onions
- Let y to be the number of tomatoes
- $x+y=13$ (i)
- $xy=40$ (ii)

Equation (i) and (ii) can be solved simultaneously, i.e.

$y^2-13y+40=0$ (iii)

(iii) is a quadratic equation which can be solved by numerous methods.
Using formula method

$$y=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad (2)$$

With a=1, b= -13 and c=40, we obtain the solution as $y=8$ or $y=5$. Then $x=5$ or $x=8$
Therefore the numbers of onions are either 8 or 5, also the number of tomatoes are either 8 or 5

Analysis

- double the sum of tomatoes and onions is 26, i.e 2(8+5)=26
- Double the product of onions and tomatoes is 80, i.e 2(8*5)=80

We observe that any of the tomatoes and onions can take the value of 8 or 5 respectively; this is because of commutative properties of addition and multiplication of real numbers.

4. Conclusion

This paper provides evidence that mathematics as a universal language was not necessarily introduced into Hausa culture (the dominant culture of northern Nigeria) as a result of Western education because mathematical game in Hausa culture existed before the advent of Western education. Elders in Hausa society used to give Hausa game to the younger ones with hope of shaping their intellectual thinking. The existence of RAS KONA UWAR LISSAFI (idiot savants who are the fastest answer-givers to Hausa game) phenomenon in Hausa culture underscores the practice of Hausa game.Given the foregoing, and the highlighted applications of mathematics in Hausa culture, it is hoped that experts from different fields such as education, mathematics, psychology, and game design will put heads together to explore and expand the empirical studies in the use of mathematics in Hausa culture with the possibilities of extending the work to other areas. This might even encourage our younger ones who dread mathematics to have more interest in the subject.

5. References

Ascher, M.(1991): Ethnomathematics: A Multicultural View of Mathematical *Ideas*. Pacific Grove: Brooks/Cole.

Auwalu, G. R.(2003):Wasakwakwalwa CTV Kano Nigeria.

Glorin, G. (1980): Connecting mathematics practices in and out of Schools, Vol. 3, No. 2 journal of Ethnomathematics Canada.

Habibu, B. (2005): Ethnomathematics: The case of mallam Na Bakin Kogi, Bayero University, Kano.Nigeria.

Heather,C .(2003) :*Success with Mathematics*" Routledge taylor and Francis Group, New York.

James , R. (1982): How students can own mathematics" Vol. 5, No. 1journal of Ethnomathematics

John,M. (1998):Ethnomathematics concept, Definition and Research perspectives, ISGEm Newsletter vol. II No. 1 Dec 1998.

Kahn,P.(2001), Studying Mathematics and its Applications, Basingstoke, palgrave.

Louis, F. (1986): Candy Selling and math learning, ISGEm Newsletter vol. 4 No. 2

Mohammed,W.Y. and et al (2010) : Ethnomathematics A case of Wasakwakwalwa (Hausa culture puzzles) in Northern Nigeria, International Journal of Basic and Applied Sciences,vol:10 No:1,pp:14-24.

Musa, H. I (2007): Wasakwakwalwa" Radio FM Kano, unpublished

Tandarki, W. and et al (2007): Wasakwakwalwa Freedom FM Radio, Kano, Nigeria.

Yahaya, I. Y. (1977) :Labarun Gargajiya *1*, Ibadan: Oxford University Press.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 3, No. 1, pp 43 – 52, 2010

A COEFFICIENT INEQUALITY FOR A SUB-CLASS OF ANALYTIC FUNCTIONS

B.S. MEHROK
 # 643 E, B.R.S. Nagar,
Ludhiana (Punjab), India

DEEPAK GUPTA
Department of Mathematics
M.M.University , Mullana-Ambala (Haryana), India
E-Mail : guptadeepak@yahoo.co.in

GAGANDEEP SINGH
Department of Mathematics
Rayat Polytechnic College , Railmajra (Punjab), India
E-Mail : kamboj.gagandeep@ yahoo.in

Abstract . In this paper a coefficient inequality for a subclass of analytic functions is obtained. Results proved by various authors will follow as special cases from our theorem.

Mathematics Subject Classification : 30C45.

Keywords: Subordination, Convolution, Univalent functions, Star like functions, Convex functions.

1. Introduction

Let U be the class of bounded functions

$$w(z) = \sum_{k=1}^{\infty} c_k z^k \tag{1.1}$$

which are regular in E and satisfying the conditions $w(0) = 0$ and $|w(z)| < 1$.

It is well known [6] that $|c_1| \le 1$, $|c_2| \le 1 - |c_1|^2$.

Let A denote the class of functions

$$f(z) = z + \sum_{k=2}^{\infty} a_k z^k \quad \text{regular in } E. \tag{1.2}$$

Let $S^*(\beta)$ be the class of starlike functions of order β of the form (1.2) and with the condition

$$\mathrm{Re}\left\{\frac{zf'(z)}{f(z)}\right\} > \beta, \text{ for some } \ 0 \le \beta < 1, \ \ z \in E. \tag{1.3}$$

$K(\beta)$ is the class of convex functions of order β of the form (1.2) with the condition

$$\operatorname{Re}\left\{1+\frac{zf''(z)}{f'(z)}\right\} > \beta, 0 \le \beta < 1, z \in E. \tag{1.4}$$

Obviously $S^*(0) \equiv S$ and $K(0) \equiv K$.

The classes $S^*(\beta)$ and $K(\beta)$ were introduced by Robertson[7].

A function $f(z) \in A$ is called α - convex function of order β ($0 \le \beta < 1$) if it satisfies the following conditions

(i) $\dfrac{f(z).f'(z)}{z} \neq 0, \quad z \in E.$

(ii) $\operatorname{Re} J(\alpha; f) > \beta$ where

$$J(\alpha; f) = (1-\alpha)\frac{zf'(z)}{f(z)} + \alpha\left(1+\frac{zf''(z)}{f'(z)}\right), z \in E, \alpha \text{ any real number.} \tag{1.5}$$

The class of α -convex functions of order β is denoted by $M(\alpha, \beta)$.

Clearly $M(\alpha,0) \equiv M(\alpha)$, the class introduced by Mocanu[5].

It is known [4] that all α -convex functions are α -starlike in E.

A function $f(z)$ is said to be subordinate to function $g(z)$,if there exists a function $w(z) \in U$ such that $f(z) = g(w(z))$ and we write as $f(z) \prec g(z)$.

Let $M(\alpha; A, B; \delta)$ be the class of functions $f(z) \in A$ and with the conditions

(i) $\dfrac{f(z).f'(z)}{z} \neq 0$ and for $\alpha \ge 0$

(ii) $(1-\alpha)\dfrac{zf'(z)}{f(z)} + \alpha\left(1+\dfrac{zf''(z)}{f'(z)}\right) \prec \left(\dfrac{1+Az}{1+Bz}\right)^{\delta}$, $-1 \le B < A \le 1$, $0 < \delta \le 1$.

By A_p, we denote the class of functions

$$f(z) = z^p + \sum_{k=1}^{\infty} a_{p+k} z^{p+k} \ (p \text{ is a positive integer}) \text{ regular in } E. \tag{1.6}$$

If $f(z) = z^p + \sum_{k=1}^{\infty} a_{p+k} z^{p+k}$ and $g(z) = z^p + \sum_{k=1}^{\infty} b_{p+k} z^{p+k}$ are in A_p, then convolution or Hadamard product of $f(z)$ and $g(z)$ is denoted by $f * g$ and defined as

$$(f * g)(z) = z^p + \sum_{k=1}^{\infty} a_{p+k} b_{p+k} z^{p+k}.$$

Let $T_{n,p}(\alpha; A, B; \delta)$ be the class of functions $f(z) \in A_p$ and satisfying the conditions

$$\frac{f(z).f'(z)}{z^{2p-1}} \neq 0 \qquad \text{and}$$

$$J_{n,p}(\alpha; f) = (1-\alpha)\frac{z\left(D_{n-1}^{p} f(z)\right)'}{D_{n-1}^{p} f(z)} + \alpha \frac{z\left(D_{n}^{p} f(z)\right)'}{D_{n}^{p} f(z)} \prec p\left(\frac{1+Az}{1+Bz}\right)^{\delta}, \qquad (1.7)$$

where n is any integer greater than –*p*, $-1 \le B < A \le 1$, $0 < \delta \le 1$, $\alpha \ge 0$,

$$D_{n-1}^{p} f(z) = \frac{z^{p}}{(1-z)^{n+p}} * f(z).$$

$*$ stands for convolution of functions.

By $S^{*}(A, B; \delta)$,we denote the class of functions $f(z) \in A$ regular in *E* with the condition

$$\frac{zf'(z)}{f(z)} \prec \left(\frac{1+Az}{1+Bz}\right)^{\delta}, -1 \le B < A \le 1, 0 < \delta \le 1. \qquad (1.8)$$

Obviously $S^{*}(A, B;1) \equiv S^{*}(A, B)$,which is a subclass of starlike functions studied by Goel and Mehrok [1] .

Let $K(A, B; \delta)$ be the class of functions $f(z) \in A$ and satisfying the condition

$$1 + \frac{zf''(z)}{f'(z)} \prec \left(\frac{1+Az}{1+Bz}\right)^{\delta}, -1 \le B < A \le 1, 0 < \delta \le 1 \qquad (1.9)$$

Clearly $K(A, B;1) \equiv K(A, B)$, which is a subclass of Convex functions studied by Goel and Mehrok [1].

We have the following observations:

(i) $T_{0,1}(\alpha; A, B; \delta) \equiv M(\alpha; A, B; \delta)$.

(ii) $T_{0,1}(\alpha; A, B;1) \equiv M(\alpha; A, B)$.

(iii) $T_{0,1}(0; A, B; \delta) \equiv S^{*}(A, B; \delta)$.

(iv) $T_{0,1}(0;1,-1;1) \equiv S^{*}$.

(v) $T_{0,1}(0;1-2\beta,-1;1) \equiv S^{*}(\beta)$.

(vi) $T_{0,1}(1; A, B; \delta) \equiv K(A, B; \delta)$.

(vii) $T_{0,1}(1;1,-1;1) \equiv K$.

In this paper, we obtain sharp bounds of $\left|a_{p+2} - \mu a_{p+1}^{2}\right|$ for the class $T_{n,p}(\alpha; A, B; \delta)$.

Results due to Keogh and Merkes [3], Szynal [8] and Goel and Mehrok [2] follows as special cases from our theorem.

2. Main Result

Theorem 2.1 If $f \in T_{n,p}(\alpha; A, B; \delta)$, then

(i) for μ complex ,

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \leq \begin{cases} \dfrac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}; |\lambda - \mu| \leq \gamma \\ \dfrac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}|\lambda - \mu|; |\lambda - \mu| \geq \gamma \end{cases} \tag{2.1}$$

(ii) for μ real ,

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \leq \begin{cases} \dfrac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}(\lambda - \mu); \mu \leq \lambda - \gamma \\ \dfrac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}; \lambda - \gamma \leq \mu \leq \lambda + \gamma \\ \dfrac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}(\mu - \lambda); \mu \geq \lambda + \gamma \end{cases} \tag{2.2}$$

where

$$\lambda = \frac{p\delta(A-B)\left[(n+p)^2 + \alpha(2n+2p+1)\right] + (n+p+\alpha)^2\left[\dfrac{\delta(A-B)}{2} - \dfrac{(A+B)}{2}\right]}{p\delta(A-B)(n+p+1)(n+p+2\alpha)} \tag{2.3}$$

and $$\gamma = \frac{(n+p+\alpha)^2}{p\delta(A-B)(n+p+1)(n+p+2\alpha)}. \tag{2.4}$$

Proof. As $f \in T_{n,p}(\alpha; A, B; \delta)$, therefore

$$J_{n,p}(\alpha; f) = (1-\alpha)\frac{z\left(D_{n-1}^p f(z)\right)'}{D_{n-1}^p f(z)} + \alpha\frac{z\left(D_n^p f(z)\right)'}{D_n^p f(z)} \prec p\left(\frac{1+Az}{1+Bz}\right)^{\delta}.$$

By definition of subordination,

$$(1-\alpha)\frac{z\left(D_{n-1}^p f(z)\right)'}{D_{n-1}^p f(z)} + \alpha\frac{z\left(D_n^p f(z)\right)'}{D_n^p f(z)} = p\left(\frac{1+Aw(z)}{1+Bw(z)}\right)^{\delta} \text{ ,where } w(z) \in U. \tag{2.5}$$

An easy calculation shows that

$$\frac{z\left(D_{n-1}^{p}f(z)\right)'}{D_{n-1}^{p}f(z)} = p+(n+p)a_{p+1}z+(n+p)\left[(n+p+1)a_{p+2}-(n+p)a_{p+1}^{2}\right]z^{2}+\ldots.. \quad (2.6)$$

Replacing n by n+1 in (2.6) ,we have

$$\frac{z\left(D_{n}^{p}f(z)\right)'}{D_{n}^{p}f(z)} = p+(n+p+1)a_{p+1}z+(n+p+1)\left[(n+p+2)a_{p+2}-(n+p+1)a_{p+1}^{2}\right]z^{2}+\ldots.. \quad (2.7)$$

Using (2,6) and (2.7) , (2.5) becomes

$$p+(n+p+\alpha)a_{p+1}z+\left\{(n+p+1)(n+p+2\alpha)a_{p+2}-[(n+p)^{2}+\alpha(2n+2p+1)]a_{p+1}^{2}\right\}z^{2}+\ldots...$$

$$= p+p\delta(A-B)c_{1}z+\left[p\delta(A-B)\left(c_{2}-Bc_{1}^{2}\right)+\frac{p\delta(\delta-1)}{2}(A-B)^{2}c_{1}^{2}\right]z^{2}+\ldots.. \quad (2.8)$$

On equating coefficients in (2.8) , we get

$$a_{p+1} = \frac{p\delta(A-B)}{(n+p+\alpha)}c_{1} \quad (2.9)$$

and

$$a_{p+2} = \frac{p\delta(A-B)\left(c_{2}-Bc_{1}^{2}\right)+\frac{p\delta(\delta-1)}{2}(A-B)^{2}c_{1}^{2}+\left[(n+p)^{2}+\alpha(2n+2p+1)\right]\frac{p^{2}\delta^{2}(A-B)^{2}c_{1}^{2}}{(n+p+\alpha)^{2}}}{(n+p+1)(n+p+2\alpha)} \quad (2.10)$$

From (2.9) and (2.10) ,we have

$$a_{p+2}-\mu a_{p+1}^{2} = \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}c_{2}$$

$$+\frac{p^{2}\delta^{2}(A-B)^{2}}{(n+p+\alpha)^{2}}\left\{\frac{p\delta(A-B)\left[(n+p)^{2}+\alpha(2n+2p+1)\right]+(n+p+\alpha)^{2}\left[\frac{\delta(A-B)}{2}-\frac{(A+B)}{2}\right]}{p\delta(A-B)(n+p+1)(n+p+2\alpha)}-\mu\right\}c_{1}^{2}.$$

$$\text{So}\left|a_{p+2}-\mu a_{p+1}^{2}\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}\left|c_{2}\right|+\frac{p^{2}\delta^{2}(A-B)^{2}}{(n+p+\alpha)^{2}}\left|\lambda-\mu\right|c_{1}^{2}\Big|. \quad (2.11)$$

By using $|c_2| \le 1 - |c_1|^2$, (2.11) becomes

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)} + \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}\left[|\lambda-\mu| - \gamma\right]|c_1|^2. \qquad (2.12)$$

If $|\lambda - \mu| \le \gamma$, then from (2.12)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}.$$

If $|\lambda - \mu| \ge \gamma$, then again from (2.12)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}|\lambda-\mu|.$$

Now we discuss the case when μ is real .

Case I. for $\mu \le \lambda$, from (2.12)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)} + \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}\left[(\lambda-\gamma) - \mu\right]|c_1|^2. \qquad (2.13)$$

If $\mu \le \lambda - \gamma$, then (2.13) becomes

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}(\lambda-\mu).$$

If $\lambda - \gamma \le \mu \le \lambda$, from (2.13)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}.$$

Case II. for $\mu \ge \lambda$, from (2.12)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)} + \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}\left[\mu - (\lambda+\gamma)\right]|c_1|^2. \qquad (2.14)$$

If $\mu \le \lambda + \gamma$, then from (2.14)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p\delta(A-B)}{(n+p+1)(n+p+2\alpha)}.$$

If $\mu \ge \lambda + \gamma$, from (2.14)

$$\left|a_{p+2} - \mu a_{p+1}^2\right| \le \frac{p^2\delta^2(A-B)^2}{(n+p+\alpha)^2}(\mu-\lambda).$$

On putting $n = 0$, $p = 1$, $\alpha = 0$ in the above theorem , we have the following:

Cor 2.1 If $f \in S^*(A,B;\delta)$, then

(i) for μ complex,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{\delta(A-B)}{2}; \left|\lambda_1-\mu\right| \le \gamma_1 \\ \delta^2(A-B)^2\left|\lambda_1-\mu\right|; \left|\lambda_1-\mu\right| \ge \gamma_1 \end{cases}$$

(ii) for μ real,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \delta^2(A-B)^2(\lambda_1-\mu); \mu \le \lambda_1-\gamma_1 \\ \dfrac{\delta(A-B)}{2}; \lambda_1-\gamma_1 \le \mu \le \lambda_1+\gamma_1 \\ \delta^2(A-B)^2(\mu-\lambda_1); \mu \ge \lambda_1+\gamma_1 \end{cases}$$

where $\lambda_1 = \dfrac{(A-B)(3\delta-1)-2B}{4\delta(A-B)}$ and $\gamma_1 = \dfrac{1}{2\delta(A-B)}$.

If we put $n=0$, $p=1$, $\alpha=1$ in theorem (2.1), we get the following result:

Cor 2.2 If $f \in K(A,B;\delta)$, then

(i) for μ complex,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{\delta(A-B)}{6}; \left|\lambda_2-\mu\right| \le \gamma_2 \\ \dfrac{\delta^2(A-B)^2}{4}\left|\lambda_2-\mu\right|; \left|\lambda_2-\mu\right| \ge \gamma_2 \end{cases}$$

(ii) for μ real,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{\delta^2(A-B)^2}{4}(\lambda_2-\mu); \mu \le \lambda_2-\gamma_2 \\ \dfrac{\delta(A-B)}{6}; \lambda_2-\gamma_2 \le \mu \le \lambda_2+\gamma_2 \\ \dfrac{\delta^2(A-B)^2}{4}(\mu-\lambda_2); \mu \ge \lambda_2+\gamma_2 \end{cases}$$

where $\lambda_2 = \dfrac{(A-B)(3\delta-1)-2B}{3\delta(A-B)}$ and $\gamma_2 = \dfrac{2}{3\delta(A-B)}$.

By putting $n = 0$, $p = 1$, $\alpha = 0$, $A = 1$, $B = -1$, $\delta = 1$ in theorem (2.1), we have the following result due to Keogh and Merkes [3].

Cor 2.3 If $f \in S^*$, then

$$\left|a_3 - \mu a_2^2\right| \leq \max\{1, |4\mu - 3|\}.$$

For $n = 0$, $p = 1$, $\alpha = 1$, $A = 1$, $B = -1$, $\delta = 1$, the following result is obvious:

Cor 2.4 If $f \in K$, then

$\left|a_3 - \mu a_2^2\right| \leq \max\left\{\frac{1}{3}, |1 - \mu|\right\}$ **,** which is also a result due to Keogh and Merkes [3].

If we take $n = 0$, $p = 1$, $A = 1 - 2\beta$, $B = -1$, $\delta = 1$ in the theorem (2.1), we obtain the following result due to Szynal [8].

Cor 2.5 For $f \in S^*(\beta)$, we have

$$\left|a_3 - \mu a_2^2\right| \leq \frac{1-\beta}{1+2\alpha} \max\left\{1, \frac{\left|4\mu(1-\beta)(1+2\alpha) + 4(1-\beta) - (1+\alpha)(7+\alpha-6\beta)\right|}{(1+\alpha)^2}\right\}.$$

If we put $n = 0$, $p = 1$ in theorem (2.1), we get the following result:

Cor 2.6 If $f \in M(\alpha; A, B; \delta)$, then

(i) for μ complex,

$$\left|a_3 - \mu a_2^2\right| \leq \begin{cases} \dfrac{\delta(A-B)}{2(1+2\alpha)}; |\lambda_3 - \mu| \leq \gamma_3 \\ \dfrac{\delta^2(A-B)^2}{(1+\alpha)^2}|\lambda_3 - \mu|; |\lambda_3 - \mu| \geq \gamma_3 \end{cases}$$

(ii) for μ real,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{\delta^2(A-B)^2}{(1+\alpha)^2}(\lambda_3-\mu); \mu \le \lambda_3-\gamma_3 \\ \dfrac{\delta(A-B)}{2(1+2\alpha)}; \lambda_3-\gamma_3 \le \mu \le \lambda_3+\gamma_3 \\ \dfrac{\delta^2(A-B)^2}{(1+\alpha)^2}(\mu-\lambda_3); \mu \ge \lambda_3+\gamma_3 \end{cases}$$

where $\lambda_3=\dfrac{\delta(A-B)[1+3\alpha]+(1+\alpha)^2\left[\frac{\delta(A-B)}{2}-\frac{(A+B)}{2}\right]}{2\delta(A-B)(1+2\alpha)}$ and $\gamma_3=\dfrac{(1+\alpha)^2}{2\delta(A-B)(1+2\alpha)}$.

Putting $n=0$, $p=1$, $\delta=1$ in theorem (2.1) , we have

Cor 2.7 If $f \in M(\alpha; A, B)$, then

(i) for μ complex ,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{(A-B)}{2(1+2\alpha)}; \left|\lambda_4-\mu\right| \le \gamma_4 \\ \dfrac{(A-B)^2}{(1+\alpha)^2}\left|\lambda_4-\mu\right|; \left|\lambda_4-\mu\right| \ge \gamma_4 \end{cases}$$

(ii) for μ real ,

$$\left|a_3-\mu a_2^2\right| \le \begin{cases} \dfrac{(A-B)^2}{(1+\alpha)^2}(\lambda_4-\mu); \mu \le \lambda_4-\gamma_4 \\ \dfrac{(A-B)}{2(1+2\alpha)}; \lambda_4-\gamma_4 \le \mu \le \lambda_4+\gamma_4 \\ \dfrac{(A-B)^2}{(1+\alpha)^2}(\mu-\lambda_4); \mu \ge \lambda_4+\gamma_4 \end{cases}$$

where $\lambda_4=\dfrac{(A-B)[1+3\alpha]-B(1+\alpha)^2}{2(A-B)(1+2\alpha)}$ and $\gamma_4=\dfrac{(1+\alpha)^2}{2(A-B)(1+2\alpha)}$.

These results were due to Goel and Mehrok [2].

References

[1]. Goel,R.M.and Mehrok ,B.S. (1981) ,"On the coefficients of a subclass of starlike functions ," *Ind. J. Pure. App. Math.* , 12(5) , 634-647 .
[2]. Goel, R.M. and Mehrok ,Beant Singh (1991) , " A coefficient inequality for certain classes of analytic functions ," *Tamkang Journal Of Mathematics* , 22, 153-163
[3]. Keogh , E.R. and Merkes, E.P. (1969) , "A coefficient inequality for certain classes of analytic functions ," *Proc.Amer.Math.Soc*. 20, 8-12.
[4]. Miller, S.S., Mocanu , P.T. and Reade, M.O. (1973) " All α - convex functions are univalent and Starlike," *Proc.Amer.Math.Soc*. 37, 553-554.
[5]. Mocanu , P.T. (1969), " Une propriete' de convexite' généralisée dans la théorie de la représentation conforme," *Mathematica* (CLUJ) ,11(34), 127-133.
[6]. Nehari , Z. (1952) , " Conformal mapping," *Mc.Graw Hill, New york* .
[7]. Robertson , M.S. (1936) , " On the theory of univalent functions," *Ann. Math*. 37, 374-408.
[8]. Szynal, J. (1972), "Some remarks on coefficient inequality for α convex functions," *Bull. De L'Acad. Des Sci Ser des Sci Math. Astr et phys*. Vol. XX, No. 11, 917-919.

ISSN 0974 – 3340

SUTRA

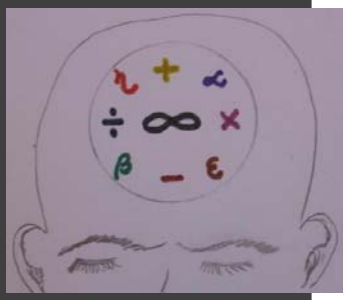
International Journal of Mathematical Science Education

VOLUME 4, ISSUE I JANUARY-JUNE 2011

Published by

Technomathematics Research Foundation, India

Sutra: International Journal of Mathematical Science Education



EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India.*

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain.*
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK.*
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India.*
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India.*
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India.*
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India.*
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India.*
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia.*
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

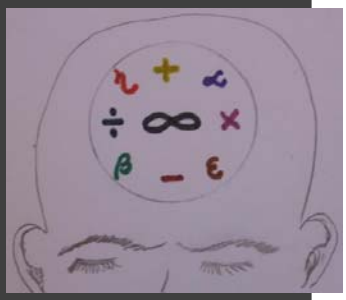
Copyright ©2011 ***Technomthematics Research Foundation.***

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.

Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.

Sutra: International Journal of Mathematical Science Education (ISSN 0974 – 3340) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education



CONTENTS (Vol. 4, No. 1, 2011)

1. **NECESSARY AND SUFFICIENT OPTIMALITY CONDITIONS FOR NONLINEAR FUZZY OPTIMIZATION PROBLEM** 1—16
V. D. PATHAK · U. M. PIRZADA

2. **SUBCLASSES OF P-VALENT STARLIKE FUNCTIONS DEFINED BY USING CERTAIN FRACTIONAL DERIVA TIVE OPERATOR** 17—32
SOMIA MUFTAH AMSHERI AND VALENTINA ZHARKOVA

3. **ANALYSIS OF THE ROLE OF ENTROPIES AND ASSIGMENT OF RANKS TO THE FEATURES IN GENRE DISCRIMINATION** 33—48
HEMLATA PANDEA AND H. S. DHAMIB

4. **IDENTIFYING AND SEQUENCING OF THE ELEMENTARY CONCEPTS OF MEASUREMENT OF MASS** 49—58
S K. SAMSUL ALAM

5. **THE CENTROID OF SOME GENERALIZED PEDAL CONFIGURATIONS** 59—72
SZILÁRD ANDRÁS

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 1, pp. 1 - 16, 2011

Necessary and Sufficient Optimality Conditions for Nonlinear Fuzzy Optimization Problem

V. D. Pathak · U. M. Pirzada

In this paper we derive the necessary and sufficient Kuhn-Tucker like optimality conditions for nonlinear fuzzy optimization problems with fuzzy valued objective function and fuzzy-valued constraints using the concept of convexity and H-diffrentiability of fuzzy-valued functions.

Keywords: Fuzzy numbers · Hukuhara differentiability · Kuhn-Tucker optimality conditions

1 Introduction

Classical optimization techniques have been successfully applied for years. In real process optimization, there exist different types of uncertainties in the system. Zimmermann [16] pointed out various kinds of uncertainties that can be categorized as stochastic uncertainty and fuzziness. The optimization under a fuzzy environment or which involve fuzziness is called fuzzy optimization.

Bellman and Zadeh in 1970 [1] proposed the concept of fuzzy decision and the decision model under fuzzy environments. After that, various approaches to fuzzy linear and nonlinear optimization, have been developed over the years by researchers.

The nondominated solution of a nonlinear optimization problem with fuzzy-valued objective function was proposed by Wu [10]. Using the concept of continuous differentiability of fuzzy-valued functions, he derived the sufficient optimality conditions for obtaining the nondominated solution of fuzzy optimization problem having fuzzy- valued objective function with real constraints. However, the fuzzy optimization problem having fuzzy-valued constraints can not be solved by using the results of Wu [10]. In this article, we establish Kuhn-Tucker like both necessary and sufficient optimality conditions for obtaining the nondominated solution of a nonlinear fuzzy optimization

V. D. Pathak
Charotar Institute of Computer Applications, Changa-388421, DI. Anand, Gujarat, India. E-mail: vdpathak@yahoo.com

U. M. Pirzada
Department of Applied Mathematics, Faculty of Tech. & Engg.,M.S.University of Baroda, Vadodara 390001, India.
E-mail: salmapirzada@yahoo.com

problem with fuzzy-valued objective function and fuzzy-valued constraints.
In Section 2, we introduce definition of fuzzy number, basic properites and arithmetics of fuzzy numbers. In Section 3, we consider the differential calculus of fuzzy-valued functions defined on $\mathbb{R}$ and $\mathbb{R}^n$ using hukuhara differentiability of fuzzy-valued functions. In Section 4, we provide nondominated solution of unconstrained fuzzy optimization problems by proving the first and second order optimality conditions. In Section 5, we provide nondominated solution of nonlinear constrained fuzzy optimization problems by proving the Kuhn-Tucker like optimality conditions for the same. And at last we conclude in Section 6.

2 Preliminaries

Definition 1 [6] Let $\mathbb{R}$ be the set of real numbers and $\tilde{a} : \mathbb{R} \to [0,1]$ be a fuzzy set. We say that $\tilde{a}$ is a fuzzy number if it satisfies the following properties:

(i) $\tilde{a}$ is normal, that is, there exists $x_0 \in \mathbb{R}$ such that $\tilde{a}(x_0) = 1$;
(ii) $\tilde{a}$ is fuzzy convex, that is, $\tilde{a}(tx + (1-t)y) \geq min\{\tilde{a}(x), \tilde{a}(y)\}$, whenever x, y $\in \mathbb{R}$ *and* $t \in [0,1]$;
(iii) $\tilde{a}(x)$ is upper semicontinuous on $\mathbb{R}$, that is, $\{x/\tilde{a}(x) \geq \alpha\}$ is a closed subset of $\mathbb{R}$ for each $\alpha \in (0,1]$;
(iv) $cl\{x \in \mathbb{R}/\tilde{a}(x) > 0\}$ forms a compact set.

The set of all fuzzy numbers on $\mathbb{R}$ is denoted by $F(\mathbb{R})$. For all $\alpha \in (0,1]$, α-level set $\tilde{a}_\alpha$ of any $\tilde{a} \in F(\mathbb{R})$ is defined as $\tilde{a}_\alpha = \{x \in \mathbb{R}/\tilde{a}(x) \geq \alpha\}$. The 0-level set $\tilde{a}_0$ is defined as the closure of the set $\{x \in \mathbb{R}/\tilde{a}(x) > 0\}$. By definition of fuzzy numbers, we can prove that, for any $\tilde{a} \in F(\mathbb{R})$ and for each $\alpha \in (0,1]$, $\tilde{a}_\alpha$ is compact convex subset of $\mathbb{R}$, and we write $\tilde{a}_\alpha = [\tilde{a}_\alpha^L, \tilde{a}_\alpha^U]$. $\tilde{a} \in F(\mathbb{R})$ can be recovered from its α-cuts by a well-known decomposition theorem (ref. [7]), which states that $\tilde{a} = \cup_{\alpha \in [0,1]} \alpha \cdot \tilde{a}_\alpha$ where union on the right-hand side is the standard fuzzy union.

Definition 2 [15] According to Zadeh's extension principle, we have addition and scalar multiplication in fuzzy number space $F(\mathbb{R})$ by their α-cuts are as follows:

$$(\tilde{a} \oplus \tilde{b})_\alpha = [\tilde{a}_\alpha^L + \tilde{b}_\alpha^L, \tilde{a}_\alpha^U + \tilde{b}_\alpha^U]$$
$$(\lambda \odot \tilde{a})_\alpha = [\lambda \cdot \tilde{a}_\alpha^L, \lambda \cdot \tilde{a}_\alpha^U],$$

where $\tilde{a}, \tilde{b} \in F(\mathbb{R})$, $\lambda \in \mathbb{R}$ and $\alpha \in [0,1]$.

Definition 3 [9] Let $A, B \subseteq \mathbb{R}^n$. The Hausdorff metric d_H is defined by

$$d_H(A,B) = \max\{\sup_{x \in A} \inf_{y \in B} ||x-y||, \sup_{y \in B} \inf_{x \in A} ||x-y||\}.$$

Then the **metric** d_F on $F(\mathbb{R})$ is defined as

$$d_F(\tilde{a}, \tilde{b}) = \sup_{0 \leq \alpha \leq 1} \{d_H(\tilde{a}_\alpha, \tilde{b}_\alpha)\},$$

for all $\tilde{a}, \tilde{b} \in F(\mathbb{R})$. Since $\tilde{a}_\alpha$ and $\tilde{b}_\alpha$ are closed bounded intervals in $\mathbb{R}$,

$$d_F(\tilde{a}, \tilde{b}) = \sup_{0 \leq \alpha \leq 1} \max\{|\tilde{a}_\alpha^L - \tilde{b}_\alpha^L|, |\tilde{a}_\alpha^U - \tilde{b}_\alpha^U|\}.$$

We need the following proposition.

Proposition 1 *[3] For $\tilde{a} \in F(\mathbb{R})$, we have*

(i) $\tilde{a}_\alpha^L$ is bounded left continuous nondecreasing function on (0,1];
(ii) $\tilde{a}_\alpha^U$ is bounded left continuous nonincreasing function on (0,1];
(iii) $\tilde{a}_\alpha^L$ and $\tilde{a}_\alpha^U$ are right continuous at $\alpha = 0$;
(iv) $\tilde{a}_\alpha^L \leq \tilde{a}_\alpha^U$.

Moreover, if the pair of functions $\tilde{a}_\alpha^L$ and $\tilde{a}_\alpha^U$ satisfy the conditions (i)-(iv), then there exists a unique $\tilde{a} \in F(\mathbb{R})$ such that $\tilde{a}_\alpha = [\tilde{a}_\alpha^L, \tilde{a}_\alpha^U]$, for each $\alpha \in [0,1]$.

We define here a partial order relation on fuzzy number space.

Definition 4 For $\tilde{a}$ and $\tilde{b}$ in $F(\mathbb{R})$ and $\tilde{a}_\alpha = [\tilde{a}_\alpha^L, \tilde{a}_\alpha^U]$ and $\tilde{b}_\alpha = [\tilde{b}_\alpha^L, \tilde{b}_\alpha^U]$ are two closed intervals in $\mathbb{R}$, for all $\alpha \in [0,1]$, we define

(i) $\tilde{a} \preceq \tilde{b}$ if and only if $\tilde{a}_\alpha^L \leq \tilde{b}_\alpha^L$ and $\tilde{a}_\alpha^U \leq \tilde{b}_\alpha^U$ for all $\alpha \in [0,1]$;
(ii) $\tilde{a} \prec \tilde{b}$ if and only if

$$\begin{cases} \tilde{a}_\alpha^L < \tilde{b}_\alpha^L \\ \tilde{a}_\alpha^U \leq \tilde{b}_\alpha^U \end{cases} \text{ or } \begin{cases} \tilde{a}_\alpha^L \leq \tilde{b}_\alpha^L \\ \tilde{a}_\alpha^U < \tilde{b}_\alpha^U \end{cases} \text{ or } \begin{cases} \tilde{a}_\alpha^L < \tilde{b}_\alpha^L \\ \tilde{a}_\alpha^U < \tilde{b}_\alpha^U \end{cases}$$

for all $\alpha \in [0,1]$.

"$\preceq$" is partial order relation on fuzzy number space.

Definition 5 [14] The membership function of a triangular fuzzy number $\tilde{a}$ is defined by

$$\zeta_{\tilde{a}}(r) = \begin{cases} \frac{(r-a^L)}{(a-a^L)} & if \quad a^L \leq r \leq a \\ \frac{(a^U-r)}{(a^U-a)} & if \quad a < r \leq a^U \\ 0 & otherwise \end{cases}$$

which is denoted by $\tilde{a} = (a^L, a, a^U)$. The α-level set of $\tilde{a}$ is then

$$\tilde{a}_\alpha = [(1-\alpha)a^L + \alpha a, (1-\alpha)a^U + \alpha a].$$

3 Differential calculus of fuzzy-valued function

3.1 Continuity of fuzzy-valued function

Definition 6 [8] Let V be a real vector space and $F(\mathbb{R})$ be a fuzzy number space. Then a function $\tilde{f} : V \to F(\mathbb{R})$ is called fuzzy-valued function defined on V.
Corresponding to such a function $\tilde{f}$ and $\alpha \in [0,1]$, we define two real-valued functions $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$ on V as $\tilde{f}_\alpha^L(x) = (\tilde{f}(x))_\alpha^L$ and $\tilde{f}_\alpha^U(x) = (\tilde{f}(x))_\alpha^U$ for all $x \in V$.

Definition 7 [4] Let $\tilde{f} : \mathbb{R}^n \to F(\mathbb{R})$ be a fuzzy-valued function. We say that $\tilde{f}$ is continuous at $c \in \mathbb{R}^n$ if for every $\epsilon > 0$, there exists a $\delta = \delta(c, \epsilon) > 0$ such that

$$d_F(\tilde{f}(x), \tilde{f}(c)) < \epsilon$$

for all $x \in \mathbb{R}^n$ with $\|x - c\| < \delta$. That is,

$$\lim_{x \to c} \tilde{f}(x) = \tilde{f}(c).$$

We prove the following proposition.

Proposition 2 *Let $\tilde{f} : \mathbb{R}^n \to F(\mathbb{R})$ be a fuzzy-valued function. If $\tilde{f}$ is continuous at $c \in \mathbb{R}^n$, then functions $\tilde{f}_\alpha^L(x)$ and $\tilde{f}_\alpha^U(x)$ are continuous at c, for all $\alpha \in [0,1]$.*

Proof The result follows using the definitions of continuity of fuzzy-valued function $\tilde{f}$ and metric on fuzzy numbers. □

3.2 H-differentiability of fuzzy-valued function on $\mathbb{R}$

Let $\tilde{a}$ and $\tilde{b}$ be two fuzzy numbers. If there exists a fuzzy number $\tilde{c}$ such that $\tilde{c} \oplus \tilde{b} = \tilde{a}$. Then $\tilde{c}$ is called **Hukuhara difference** of $\tilde{a}$ and $\tilde{b}$ and is denoted by $\tilde{a} \ominus_H \tilde{b}$.

H-differentiability of fuzzy-valued function due to M.L. Puri and D.A. Ralescu [11] is as follows

Definition 8 Let X be a subset of $\mathbb{R}$. A fuzzy-valued function $\tilde{f} : X \to F(\mathbb{R})$ is said to be H-differentiable at $x^0 \in X$ if there exists a fuzzy number $D\tilde{f}(x^0)$ such that the limits (with respect to metric d_F)

$$\lim_{h \to 0^+} \frac{1}{h} \odot [\tilde{f}(x^0 + h) \ominus_H \tilde{f}(x^0)], \quad \textit{and} \quad \lim_{h \to 0^+} \frac{1}{h} \odot [\tilde{f}(x^0) \ominus_H \tilde{f}(x^0 - h)]$$

both exist and are equal to $D\tilde{f}(x^0)$. In this case, $D\tilde{f}(x^0)$ is called the H-derivative of $\tilde{f}$ at x^0. If $\tilde{f}$ is H-differentiable at any $x \in X$, we call $\tilde{f}$ is H-differentiable over X.

Remark 1 Many fuzzy-valued functions are H-differentaible for which Hukuhara differences $\tilde{f}(x^0 + h) \ominus_H \tilde{f}(x^0)$ and $\tilde{f}(x^0) \ominus_H \tilde{f}(x^0 - h)$ both exist. The following example illustrates the fact.

Example 1 Given in [11], let $\tilde{f} : (0, 2\pi) \to F(\mathbb{R})$ be defined on level sets by

$$[\tilde{f}(x)]_\alpha = (1 - \alpha)(2 + \sin(x))[-1, 1],$$

for $\alpha \in [0,1]$. At $x^0 = \pi/2$, H-difference does not exist. Therefore, function is not H-differentiable at $x^0 = \pi/2$.

Now we prove following proposition regarding differentiability of $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$.

Proposition 3 *Let X be a subset of $\mathbb{R}$. If a fuzzy-valued function $\tilde{f} : X \to F(\mathbb{R})$ is H-differentiable at x^0 with derivative $D\tilde{f}(x^0)$, then $\tilde{f}_\alpha^L(x)$ and $\tilde{f}_\alpha^U(x)$ are differentiable at x^0, for all $\alpha \in [0,1]$. Moreover, we have $(D\tilde{f})_\alpha(x^0) = [D(\tilde{f}_\alpha^L)(x^0), D(\tilde{f}_\alpha^U)(x^0)]$.*

Proof The result follows from definitions of H-differentiability of fuzzy-valued function and metric on fuzzy number space. □

3.3 H-differentiability of fuzzy-valued function on $\mathbb{R}^n$

Definition 9 [10] Let $\tilde{f}$ be a fuzzy-valued function defined on an open subset X of $\mathbb{R}^n$ and let $x^0 = (x_1^0, ..., x_n^0) \in X$ be fixed.
We say that $\tilde{f}$ has the i^{th} partial H-derivative $D_i\tilde{f}(x^0)$ at x^0 if the fuzzy-valued function $\tilde{g}(x_i) = \tilde{f}(x_1^0, .., x_{i-1}^0, x_i, x_{i+1}^0, .., x_n^0)$ is H-differentiable at x_i^0 with H-derivative $D_i\tilde{f}(x^0)$. We also write $D_i\tilde{f}(x^0)$ as $(\partial\tilde{f}/\partial x_i)(x^0)$.

Definition 10 [10] We say that $\tilde{f}$ is H-differentiable at x^0 if one of the partial H-derivatives $\partial\tilde{f}/\partial x_1, ..., \partial\tilde{f}/\partial x_n$ exists at x^0 and the remaining n-1 partial H-derivatives exist on some neighborhoods of x^0 and are continuous at x^0 (in the sense of fuzzy-valued function).

The gradient of $\tilde{f}$ at x^0 is denoted by

$$\nabla\tilde{f}(x^0) = (D_1\tilde{f}(x^0), ..., D_n\tilde{f}(x^0)),$$

and it defines a fuzzy-valued function from X to $F^n(\mathbb{R}) = F(\mathbb{R}) \times .... \times F(\mathbb{R})$ (n times), where each $D_i\tilde{f}(x^0)$ is a fuzzy number for i = 1,...,n. The α-level set of $\nabla\tilde{f}(x^0)$ is defined and denoted by

$$(\nabla\tilde{f}(x^0))_\alpha = ((D_1\tilde{f}(x^0))_\alpha, ..., (D_n\tilde{f}(x^0))_\alpha),$$

where

$$(D_i\tilde{f}(x^0))_\alpha = [D_i\tilde{f}_\alpha^L(x^0), D_i\tilde{f}_\alpha^U(x^0)],$$

i = 1,...,n.

We say that $\tilde{f}$ is H-differentiable on X if it is H-differentiable at every $x^0 \in X$.

Proposition 4 *Let X be an open subset of $\mathbb{R}^n$. If a fuzzy-valued function $\tilde{f} : X \to F(\mathbb{R})$ is H-differentiable on X. Then $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$ are also differentiable on X, for all $\alpha \in [0, 1]$. Moreover, for each $x \in X$, $(D_i\tilde{f}(x))_\alpha = [D_i\tilde{f}_\alpha^L(x), D_i\tilde{f}_\alpha^U(x)]$, i = 1,...,n.*

Proof The result follows from Propositions 2 and 3. □

Definition 11 We say that $\tilde{f}$ is continuously H-differentiable at x^0 if all of the partial H-derivatives $\partial\tilde{f}/\partial x_i$, i = 1,..,n, exist on some neighborhoods of x^0 and are continuous at x^0 (in the sense of fuzzy-valued function).
We say that $\tilde{f}$ is continuously H-differentiable on X if it is continuously H-differentiable at every $x^0 \in X$.

Proposition 5 *Let $\tilde{f} : X \to F(\mathbb{R})$ is continuously H-differentiable on X. Then $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$ are also continuously differentiable on X, for all $\alpha \in [0, 1]$.*

Proof Followed by Propositions 2 and 4. □

Now we define twice continuously H-differentiable fuzzy-valued function.

Definition 12 Let $\tilde{f} : X \to F(\mathbb{R}), X \subset \mathbb{R}^n$ be a fuzzy-valued function. Suppose now that there is $x^0 \in X$ such that gradient of $\tilde{f}$, $\nabla\tilde{f}$, is itself H-differentiable at x^0, that is, for each i, the function $D_i\tilde{f} : X \to F(\mathbb{R})$ is H-differentiable at x^0. Denote the H-partial derivative of $D_i\tilde{f}$ in the direction of $\bar{e}_j$ at x^0 by

$$D^2_{ij}\tilde{f} \;\; or \;\; \frac{\partial^2\tilde{f}(x^0)}{\partial x_i \partial x_j}, \;\; if \;\; i \neq j,$$

and

$$D^2_{ii}\tilde{f} \;\; or \;\; \frac{\partial^2\tilde{f}(x^0)}{\partial x_i^2}, \;\; if \;\; i = j.$$

Then we say that $\tilde{f}$ is twice H-differentiable at x^0, with second H-derivative $\nabla^2\tilde{f}(x^0)$ which is denoted by

$$\nabla^2\tilde{f}(x^0) = \begin{pmatrix} \frac{\partial^2\tilde{f}(x^0)}{\partial x_1^2} & \ldots & \frac{\partial^2\tilde{f}(x^0)}{\partial x_1 \partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial^2\tilde{f}(x^0)}{\partial x_n \partial x_1} & \ldots & \frac{\partial^2\tilde{f}(x^0)}{\partial x_n^2} \end{pmatrix}$$

where $\frac{\partial^2\tilde{f}(x^0)}{\partial x_i \partial x_j} \in F(\mathbb{R})$, i,j = 1,...,n.
If $\tilde{f}$ is twice H-differentiable at each x^0 in X, we say that $\tilde{f}$ is twice H-differentiable on X, and if for each i, j = 1,...,n, the cross-partial derivative $\frac{\partial^2\tilde{f}}{\partial x_i \partial x_j}$ is continuous function from X to $F(\mathbb{R})$, we say that $\tilde{f}$ is twice continuously H-differentiable on X.
The α-level set of $\nabla^2\tilde{f}(x^0)$ is defined and denoted in matrix notation as

$$(\nabla^2\tilde{f}(x^0))_\alpha = ((D^2_{ij}\tilde{f}(x^0))_\alpha)$$

i,j = 1,...,n and $\alpha \in [0,1]$, where $(D^2_{ij}\tilde{f}(x^0))_\alpha$ denotes α-cut of $(D^2_{ij}\tilde{f}(x^0))$.

Proposition 6 *Let $\tilde{f} : X \subseteq \mathbb{R}^n \to F(\mathbb{R})$ is differentiable with derivative $\nabla\tilde{f}$ on X and let each $D_i\tilde{f} : X \to F(\mathbb{R})$, i = 1,...,n, is also differentiable at x^0 with derivative $D^2_{ij}\tilde{f}(x^0)$, i , j = 1,...,n. Then $D_i\tilde{f}^L_\alpha$ and $D_i\tilde{f}^U_\alpha$ are also differentiable at x^0, for all $\alpha \in [0,1]$. Also, we have $(D^2_{ij}\tilde{f}(x^0))_\alpha = [D^2_{ij}(\tilde{f}^L_\alpha)(x^0), D^2_{ij}(\tilde{f}^U_\alpha)(x^0)]$, i,j = 1,...,n.*

Proof Follows by Proposition 5.

In order to define the Kuhn-Tucker like optimality conditions for nonlinear fuzzy optimization problems, we need to provide some properties of fuzzy-valued functions. For that first we state here following two Propositions from Real Analysis.

Proposition 7 *[13] Let ϕ be a real-valued function of two variables defined on I×[a,b], where I is an interval in $\mathbb{R}$. Suppose that the following conditions are satisfied:*

(i) For every $x \in I$, the real-valued function $h(y) = \phi(x,y)$ is Riemann integrable on [a,b]. In this case, we write $f(x) = \int_a^b \phi(x,y)\,dy$;

(ii) Let $x^0 \in int(I)$, the interior of I. For every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\left|\frac{\partial\phi}{\partial x}(x,y) - \frac{\partial\phi}{\partial x}(x^0,y)\right| < \epsilon$$

for all $y \in [a,b]$ and all $x \in (x^0 - \delta, x^0 + \delta)$.

Then $\frac{\partial \phi}{\partial x}(x^0, y)$ *is Riemann integrable on [a,b],* $f^{'}(x^0)$ *exists, and*

$$f^{'}(x^0) = \int_a^b \frac{\partial \phi}{\partial x}(x^0, y)\, dy.$$

Proposition 8 *[2] Every function monotonic on an interval is Riemann integrable there.*

Let $\tilde{f} : X \to F(\mathbb{R})$ be a fuzzy-valued function defined on X subset of $\mathbb{R}^n$. Then for each $\alpha \in [0,1]$, $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$ are real-valued functions defined on X. For any fixed $x^0 \in X$, we have the corresponding real-valued functions $\tilde{f}_\alpha^L(x^0)$ and $\tilde{f}_\alpha^U(x^0)$ defined on $\alpha \in$[0,1]. By Proposition 1 and 8, we can easily say that $\tilde{f}_\alpha^L(x^0)$ and $\tilde{f}_\alpha^U(x^0)$ are riemann integrable. So, we define new functions F^L and F^U as follows

$$F^L(x) = \int_0^1 \tilde{f}_\alpha^L(x) d\alpha \;\; and \;\; F^U(x) = \int_0^1 \tilde{f}_\alpha^U(x) d\alpha \tag{3.1}$$

for every $x \in X$. Then we have following useful proposition.

Proposition 9 *[10] Let* $\tilde{f}$ *be a fuzzy-valued function defined on an open subset X of* $\mathbb{R}^n$*. If* $\tilde{f}$ *is continuously H-differentiable on some neighborhood of* x^0*. Then the real-valued functions* F^L *and* F^U *defined as above are continuously differentiable at* x^0 *and*

$$\frac{\partial F^L}{\partial x_i}(x^0) = \int_0^1 \frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x^0)\, d\alpha \;\; and \;\; \frac{\partial F^U}{\partial x_i}(x^0) = \int_0^1 \frac{\partial \tilde{f}_\alpha^U}{\partial x_i}(x^0)\, d\alpha$$

for all i=1,..,n.

Proof We need to show that the partial derivatives $\frac{\partial F^L}{\partial x_i}$ and $\frac{\partial F^U}{\partial x_i}$ exist on some neighborhood of x^0 and are continuous at x^0 for all i =1,..,n. Since $\tilde{f}$ is continuously H-differentiable on some neighborhood of x^0. By Proposition 5, $\tilde{f}_\alpha^L$ and $\tilde{f}_\alpha^U$ are also continuously differentiable real-valued functions at x^0 for all $\alpha \in [0,1]$. Therefore, Proposition 8 say that

$$\frac{\partial F^L}{\partial x_i}(x^0) = \int_0^1 \frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x^0)\, d\alpha \;\; and \;\; \frac{\partial F^U}{\partial x_i}(x^0) = \int_0^1 \frac{\partial \tilde{f}_\alpha^U}{\partial x_i}(x^0)\, d\alpha \tag{3.2}$$

for all i=1,..,n. Since $\frac{\partial f^L}{\partial x_i}$ is continuous at x^0, that is, for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$||x - x^0|| < \delta \;\; implies \;\; \left| \frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x) - \frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x^0) \right| < \epsilon \;\; for\ all\ \alpha \in [0,1]$$

From (3.2), we have, if $||x - x^0|| < \delta$ then

$$\begin{aligned} \left| \frac{\partial F^L}{\partial x_i}(x^0) - \frac{\partial F^L}{\partial x_i}(x) \right| &= \left| \int_0^1 [\frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x^0) - \frac{\partial \tilde{f}_\alpha^U}{\partial x_i}(x)]\, d\alpha \right| \\ &\leq \int_0^1 \left| \frac{\partial \tilde{f}_\alpha^L}{\partial x_i}(x^0) - \frac{\partial \tilde{f}_\alpha^U}{\partial x_i}(x) \right| d\alpha < \epsilon, \end{aligned}$$

for all i =1,...,n. Therefore, $\frac{\partial F^L}{\partial x_i}$ is contiuous for all i = 1,...,n. Similarly we can discuss the case of $\frac{\partial F^U}{\partial x_i}$. Hence complete the proof. □

4 Neccesary and sufficient optimality conditions for unconstrained fuzzy optimization problem

4.1 Unconstrained Fuzzy Optimization Problem

Let $T \subseteq \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and $\tilde{f}$ be fuzzy-valued function defined on T. Consider the following nonlinear fuzzy optimization problem

$$(FOP1) \qquad Minimize \;\; \tilde{f}(x) = \tilde{f}(x_1, .., x_n)$$
$$Subject \;\; to \;\; x \in T$$

We define here nondominated solutions of (FOP1).

Definition 13 Let T is an open subset of $\mathbb{R}^n$.

(i) A point $x^0 \in T$ is a locally nondominated solution of (FOP1) if there exists no $x^1(\neq x^0) \in N_\epsilon(x^0) \cap T$ such that $\tilde{f}(x^1) \prec \tilde{f}(x^0)$, $N_\epsilon(x^0)$ is a ϵ-neighborhood of x^0.
(ii) A point $x^0 \in T$ is a nondominated solution of (FOP1) if there exists no $x^1(\neq x^0) \in T$ such that $\tilde{f}(x^1) \prec \tilde{f}(x^0)$.
(iii) A point $x^0 \in T$ is a locally weak nondominated solution of (FOP1) if there exists no $x^1(\neq x^0) \in N_\epsilon(x^0) \cap T$ such that $\tilde{f}(x^1) \preceq \tilde{f}(x^0)$.
(iv) A point $x^0 \in T$ is a weak nondominated solution of (FOP1) if there exists no $x^1(\neq x^0) \in T$ such that $\tilde{f}(x^1) \preceq \tilde{f}(x^0)$.

4.2 Neccessary and sufficient optimality conditions

The first and second order neccessary and sufficient optimality conditions for real unconstrained optimization problem, given in [5], are as follows.

Theorem 1 *Let T is an open subset of $\mathbb{R}^n$.*

(i) ***(FONC)*** *Let f continuously differentiable function on T. If x^* is a local minimizer of f over T, then $\nabla f(x^*) = 0$.*
(ii) ***(SONC)*** *Let f twice continuously differentiable function on T. If x^* is a local minimizer of f over T, then $\nabla^2 f(x^*)$ is positive semidefinite.*
(iii) ***(SOSC)*** *Let f twice continuously differentiable function on T. Suppose that*
1. $\nabla f(x^*) = 0$ and
2. $\nabla^2 f(x^*)$ is positive definite.
Then x^* is a strict local minimizer of f.

We prove here necessary and sufficient optimality conditions for obtaining nondominated solution of (FOP1). For that first we prove the following proposition.

Proposition 10 *If x^0 is a locally nondominated solution of (FOP1), then x^0 is also a local minimum of real-valued functions $\tilde{f}_\alpha^L(x)$ and $\tilde{f}_\alpha^U(x)$ for all $\alpha \in [0, 1]$.*

Proof We prove this result by contradiction. Assume that x^0 is not a local minimum of $\tilde{f}_\alpha^L$ or $\tilde{f}_\alpha^U$ for at least one $\alpha \in [0, 1]$. Without loss of generality, suppose that x^0 is not a local minimum of $\tilde{f}_\alpha^L$ for $\alpha_0 \in [0, 1]$. Therefore, there exists $x^1 \in N_\epsilon(x^0) \cap T$ such that

$$\tilde{f}^L_{\alpha_0}(x^1) < \tilde{f}^L_{\alpha_0}(x^0) \tag{4.1}$$

Since x^0 is a locally nondominated solution of (FOP1), there exists no $\bar{x} \in N_\epsilon(x^0) \cap T$ such that

$$\begin{cases} \tilde{f}^L_\alpha(\bar{x}) < \tilde{f}^L_\alpha(x^0) \\ \tilde{f}^U_\alpha(\bar{x}) \leq \tilde{f}^U_\alpha(x^0) \end{cases} \text{or} \begin{cases} \tilde{f}^L_\alpha(\bar{x}) \leq \tilde{f}^L_\alpha(x^0) \\ \tilde{f}^U_\alpha(\bar{x}) < \tilde{f}^U_\alpha(x^0) \end{cases} \text{or} \begin{cases} \tilde{f}^L_\alpha(\bar{x}) < \tilde{f}^L_\alpha(x^0) \\ \tilde{f}^U_\alpha(\bar{x}) < \tilde{f}^U_\alpha(x^0) \end{cases}$$

for all $\alpha \in [0,1]$. This gives contradiction to inequality (4.1). Therefore x^0 is also a minimum of real-valued functions $\tilde{f}^L_\alpha(x)$ and $\tilde{f}^U_\alpha(x)$ for all $\alpha \in [0,1]$. □

Now we prove the first-order necessary condition.

Theorem 2 *Suppose $x^0 \in T$ is a locally nondominated solution of (FOP1). Suppose also that $\tilde{f}$ is conituously H-differentiable function on T. Then*

$$\int_0^1 \nabla \tilde{f}^L_\alpha(x^0) d\alpha + \int_0^1 \nabla \tilde{f}^U_\alpha(x^0) d\alpha = 0$$

Proof The theorem can prove easily using Proposition 10 and Theorem 1 (i). □

Next, we prove first-order sufficient condition.

Theorem 3 *Let $\tilde{f}$ is twice-contiuously H-differentiable fuzzy-valued function defined on $T \subseteq \mathbb{R}^n$. If x^0 is a locally nondominated solution of (FOP1) then $\nabla^2 F(x^0)$ is positive semidefinite matrix.*

Here

$$\nabla^2 F(x^0) = \int_0^1 \nabla^2 \tilde{f}^L_\alpha(x^0) d\alpha + \int_0^1 \nabla^2 \tilde{f}^U_\alpha(x^0) d\alpha$$

Proof Since $\tilde{f}$ twice-contiuously H-differentiable fuzzy-valued function on T. By Proposition 2 and 6, $\tilde{f}^L_\alpha$ and $\tilde{f}^U_\alpha$ are also twice-contiuously H-differentiable functions on $\mathbb{R}^n$. Also, by Proposition 10, we say that $\tilde{f}^L_\alpha$ and $\tilde{f}^U_\alpha$ has local minimum at x^0. Therefore by second order neccessary condition for real unconstrained optimization stated in Theorem 1 (ii), $\nabla^2 \tilde{f}^L_\alpha(x^0)$ and $\nabla^2 \tilde{f}^U_\alpha(x^0)$ are positive semidefinite, for $\alpha \in [0,1]$.

That is, $x^T \cdot \nabla^2 \tilde{f}^L_\alpha(x^0) \cdot x \geq 0$ and $x^T \cdot \nabla^2 \tilde{f}^U_\alpha(x^0) \cdot x \geq 0$ for all $x \in T$, $x \neq 0$ and $\alpha \in [0,1]$, where x^T is transpose of x.

Therefore,

$$x^T \cdot \int_0^1 \nabla^2 \tilde{f}^L_\alpha(x^0) \cdot x \geq 0$$

and

$$x^T \cdot \int_0^1 \nabla^2 \tilde{f}^U_\alpha(x^0) \cdot x \geq 0$$

Adding these inequalities, we get

$$x^T \cdot \nabla^2 F(x^0) \cdot x \geq 0$$

for all $x \in T$, $x \neq 0$. Therefore, $\nabla^2 F(x^0)$ is positive semidefinite. □

Now, we prove second-order sufficient condition.

Theorem 4 *Let $\tilde{f}$ is twice contiuously H-differentiable function on $T \subseteq \mathbb{R}^n$. Suppose that*

1. $\nabla F(x^0)$
2. $\nabla^2 F(x^0)$ *is positive definite.*

Then, x^0 is locally weak nondominated solution of (FOP1).

Proof We prove this result using contradiction. Suppose $x^0 \in T$ is not a locally weak nondominated solution of (FOP1). Then, there exists $x^1 \in N_\epsilon(x^0) \cap T$ such that $\tilde{f}(x^1) \prec \tilde{f}(x^0)$. That is, there exists $x^1 \in N_\epsilon(x^0) \cap T$ such that

$$\begin{cases} \tilde{f}_\alpha^L(x^1) < \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^1) \leq \tilde{f}_\alpha^U(x^0) \end{cases} \text{or} \begin{cases} \tilde{f}_\alpha^L(x^1) \leq \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^1) < \tilde{f}_\alpha^U(x^0) \end{cases} \text{or} \begin{cases} \tilde{f}_\alpha^L(x^1) < \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^1) < \tilde{f}_\alpha^U(x^0) \end{cases}$$

for all $\alpha \in [0,1]$. Therefore, we have

$$F(x^1) < F(x^0), \tag{4.2}$$

where

$$F(x) = \int_0^1 \tilde{f}_\alpha^L(x)d\alpha + \int_0^1 \tilde{f}_\alpha^U(x)d\alpha.$$

As $\tilde{f}$ is twice continuously H-differentaible function, F(x) is also. Using assumption 2 and Rayleigh's inequality (refer [5]), it follows that if $d \neq 0$, then

$$0 < \lambda_{\min}(F(x^0)\|d\|^2 \leq d^T \cdot F(x^0) \cdot d.$$

By Taylor's theorem and assumption 1,

$$\begin{aligned} F(x^0 + d) - F(x^0) &= \frac{1}{2} d^T \cdot \nabla^2 F(x^0) \cdot d + O(\|d\|^2) \\ &\geq \frac{\lambda_{\min}(\nabla^2 F(x^0))}{2} \|d\|^2 + O(\|d\|^2). \end{aligned}$$

Hence, for all d such that $\|d\|$ is sufficiently small,

$$F(x^0 + d) > F(x^0)$$

This gives contradiction to inequality (4.2). Hence proved the result. □

We give one example to illustrate the above results.

Example 2 Let $\tilde{f} : \mathbb{R}^2 \to F(\mathbb{R})$ be defined by $\tilde{f}(x_1, x_2) = (0,2,4) \odot x_1^2 \oplus (0,2,4) \odot x_2^2 \odot (1,3,5)$, where $(0,2,4)$ and $(1,3,5)$ are triangular fuzzy numbers.

By first order neccessary condition: $\nabla F(x) = 0$.

Here, $\tilde{f}_\alpha^L(x_1, x_2) = 2\alpha x_1^2 + 2\alpha x_2^2 + (1 + 2\alpha)$ and

$\tilde{f}_\alpha^U(x_1, x_2) = (4 - 2\alpha)x_1^2 + (4 - 2\alpha)x_2^2 + (5 - 2\alpha)$. Therefore,

$$\nabla f_\alpha^L(x_1, x_2) = \begin{pmatrix} 4x_1\alpha \\ 4x_2\alpha \end{pmatrix}$$

and

$$\nabla f_\alpha^U(x_1,x_2)=\begin{pmatrix}2(4-2\alpha)x_1\\2(4-2\alpha)x_2\end{pmatrix}$$

Therefore,

$$\nabla F(x_1,x_2)=\begin{pmatrix}14x_1\\14x_2\end{pmatrix}$$

This implies , $x^0=(x_1,x_2)=(0,0)$.
Now to verify second order necessary and sufficient conditions, we find $\nabla^2F(x)$:

$$\nabla^2\tilde{f}_\alpha^L(x)=\begin{pmatrix}4\alpha & 0\\0 & 4\alpha\end{pmatrix}$$

and

$$\nabla^2\tilde{f}_\alpha^U(x)=\begin{pmatrix}2(4-2\alpha) & 0\\0 & 2(4-2\alpha)\end{pmatrix}$$

Therefore,

$$\nabla^2F(x)=\begin{pmatrix}14 & 0\\0 & 14\end{pmatrix}$$

which is positive definite. Therefore, by neccessary and sufficient conditions, $x^0=(0,0)$ is a nondominated solution of given problem.

5 Neccesary and sufficient optimality conditions for constrained fuzzy optimization problem

5.1 Constrained Fuzzy Optimization Problem

Let $T\subseteq\mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and $\tilde{f},\tilde{g}_j$, for j = 1,...,m be fuzzy-valued functions defined on T. Consider the following nonlinear fuzzy optimization problem

$$(FOP2)\qquad Minimize\ \ \tilde{f}(x)=\tilde{f}(x_1,..,x_n)$$

$$Subject\ \ to\ \ \tilde{g}_j(x)\preceq\tilde{0},\quad j=1,..,m,$$

where $\tilde{0}$ is a fuzzy number defined as $\tilde{0}(r)=1$ if $r=0$ and $\tilde{0}(r)=0$ if $r\neq 0$ and its level set is $\tilde{0}_\alpha=\{0\}$ for $\alpha\in[0,1]$.

Definition 14 Let $x^0\in X=\{x\in T:\tilde{g}_j(x)\preceq\tilde{0},j=1,..,m\}$. We say that x^0 is a nondominated solution of (FOP2) if there exists no $x^1(\neq x^0)\in X$ such that $\tilde{f}(x^1)\prec\tilde{f}(x^0)$. That is, x^0 is a nondominated solution of (FOP2) if there exists no $x^1(\neq x^0)\in X$ such that

$$\begin{cases}\tilde{f}_\alpha^L(x^1)<\tilde{f}_\alpha^L(x^0)\\ \tilde{f}_\alpha^U(x^1)\leq\tilde{f}_\alpha^U(x^0)\end{cases}\text{ or }\begin{cases}\tilde{f}_\alpha^L(x^1)\leq\tilde{f}_\alpha^L(x^0)\\ \tilde{f}_\alpha^U(x^1)<\tilde{f}_\alpha^U(x^0)\end{cases}\text{ or }\begin{cases}\tilde{f}_\alpha^L(x^1)<\tilde{f}_\alpha^L(x^0)\\ \tilde{f}_\alpha^U(x^1)<\tilde{f}_\alpha^U(x^0)\end{cases}$$

for all $\alpha\in[0,1]$.

5.2 Necessary and Sufficient Optimality Conditions

Let f and $g_j, j = 1, .., m$, be real-valued functions defined on $T \subset \mathbb{R}^n$. Then we consider the following optimization problem

$$(P) \qquad \begin{aligned} &Minimize \;\; f(x) = f(x_1, .., x_n) \\ &Subject \;\; to \;\; g_j(x) \leq 0, \;\; j = 1, .., m. \end{aligned}$$

The well-known Kuhn-Tucker optimality conditions for problem (P) by S. Rangarajan in [12] is stated as follows:

Theorem 5 *Let f be a convex, continuously differentiable function mapping T into* $\mathbb{R}$*, where* $T \subset \mathbb{R}^n$ *is open and convex. For j =1,..m, the constraint functions* $g_j : T \to \mathbb{R}$ *are convex, continuously differentiable functions. Suppose there is some* $x \in T$ *such that* $g_j(x) < 0$*, j =1,..,m.*
Then x^0 *is an optimal solution of propositionblem (P) over the feasible set* $\{x \in T : g_j(x) \leq 0, j = 1, .., m\}$ *if and only if there exist multipliers* $0 \leq \mu_j \in \mathbb{R}$*, j = 1,..,m, such that the Kuhn-Tucker first order conditions hold:*

(KT-1) $\nabla f(x^0) + \sum_{j=1}^{m} \mu_j \nabla g_j(x^0) = 0$*;*
(KT-2) $\mu_j \cdot g_j(x^0) = 0$ *for all j = 1,..,m.*

First we introduce the concept of convexity for fuzzy-valued functions.

Definition 15 Let T be a convex subset of $\mathbb{R}^n$ and $\tilde{f}$ be a fuzzy-valued function defined on T. We say that $\tilde{f}$ is **convex** at x^0 if

$$\tilde{f}(\lambda x^0 + (1 - \lambda)x) \preceq (\lambda \odot \tilde{f}(x^0) \oplus ((1 - \lambda) \odot \tilde{f}(x))$$

for each $\lambda \in (0, 1)$ and $x \in T$.

Proposition 11 $\tilde{f} : T \to F(\mathbb{R})$ *is convex at* x^0 *if and only if* $\tilde{f}_\alpha^L$ *and* $\tilde{f}_\alpha^U$ *are convex at* x^0*, for all* $\alpha \in [0, 1]$*.*

Proof The result can prove easily using the concepts of arithmetic operations and partial order relation of fuzzy numbers. □

We now present the Kuhn-Tucker like optimality conditions for (FOP2).

Theorem 6 *Let the fuzzy-valued objective function* $\tilde{f} : T \to F(\mathbb{R})$ *is convex and continuously H-differentiable, where* $T \subset \mathbb{R}^n$ *is open and convex. For j = 1,..,m, the fuzzy-valued constraint functions* $\tilde{g}_j : T \to F(\mathbb{R})$ *are convex and continuously H-differentiable. Let* $X = \{x \in T \subset \mathbb{R}^n : \tilde{g}_j(x) \preceq \tilde{0}, j = 1, .., m\}$ *be a feasible set of problem (FOP) and let* $x^0 \in X$*. Suppose there is some* $x \in T$ *such that* $\tilde{g}_j(x) \prec \tilde{0}$*, j =1,..,m.*
Then x^0 *is a nondominated solution of problem (FOP2) over X if and only if there exist multipliers* $0 \leq \mu_j \in \mathbb{R}$*, j = 1,..,m, such that the Kuhn-Tucker first order conditions hold:*

(FKT-1) $\displaystyle\int_0^1 \nabla \tilde{f}_\alpha^L(x^0)\, d\alpha + \int_0^1 \nabla \tilde{f}_\alpha^U(x^0)\, d\alpha + \sum_{j=1}^{m} \mu_j \nabla \tilde{g}_{j0}^U(x^0) = 0$*;*
(FKT-2) $\mu_j \cdot \tilde{g}_{j0}^U(x^0) = 0$ *for all j = 1,..,m.*

Proof **Necessary.** Define a new function,

$$F(x) = \int_0^1 \tilde{f}_\alpha^L(x)d\alpha + \int_0^1 \tilde{f}_\alpha^U(x)d\alpha. \tag{5.1}$$

Since $\tilde{f}$ is convex and continuously H-differentiable function, by Propositions 5 and 11, we say that F(x) is convex and continuously differentiable real-valued function on T. Since x^0 is a nondominated solution of (FOP2). Then there exists no $(x^1 \neq x^0) \in X$ such that

$$\begin{cases} \tilde{f}_\alpha^L(x^1) < \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^0) \leq \tilde{f}_\alpha^U(x^0) \end{cases} \text{ or } \begin{cases} \tilde{f}_\alpha^L(x^1) \leq \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^1) < \tilde{f}_\alpha^U(x^0) \end{cases} \text{ or } \begin{cases} \tilde{f}_\alpha^L(x^1) < \tilde{f}_\alpha^L(x^0) \\ \tilde{f}_\alpha^U(x^1) < \tilde{f}_\alpha^U(x^0) \end{cases}$$

for all $\alpha \in [0,1]$.

That is, there exists no $x^1(\neq x^0) \in X$ such that

$$F(x^1) < F(x^0)$$

Therefore,

$$F(x^0) \leq F(x^1)$$

Since $\tilde{g}_j$ are convex and continuously H-differentiable functions for j = 1,..,m implies $\tilde{g}_{j\alpha}^L$ and $\tilde{g}_{j\alpha}^U$ are real-valued convex and continuously differentiable functions for all $\alpha \in [0,1]$ and j =1,..,m.

By definition of partial ordering and Proposition 1, we have

$$\begin{aligned} X &= \{x \in T \subset \mathbb{R}^n : \tilde{g}_j(x) \preceq \tilde{0}, j = 1, ..., m\} \\ &= \{x \in T \subset \mathbb{R}^n : \tilde{g}_{j\alpha}^L(x) \leq 0 \; and \; \tilde{g}_{j\alpha}^U(x) \leq 0, j = 1, ..., m\} \\ &= \{x \in T \subset \mathbb{R}^n : \tilde{g}_{j\alpha}^U(x) \leq 0, j = 1, ..., m\} \\ &= \{x \in T \subset \mathbb{R}^n : \tilde{g}_{j0}^U(x) \leq 0, j = 1, ..., m\} \end{aligned}$$

Therefore, $x^0 \in X = \{x \in T \subset \mathbb{R}^n : \tilde{g}_{j0}^U(x) \leq 0, j = 1, .., m\}$ and there is some $x \in T$ such that $\tilde{g}_{j0}^U(x) < 0$, j =1,..,m. So our problem becomes an optimization problem with real objective function F(x) subject to real constraints.

Therefore, by Theorem 5, there exist multipliers $0 \leq \mu_j \in \mathbb{R}$, j = 1,..,m, such that the following kuhn-Tucker first order conditions hold:

(KT-1) $\nabla F(x^0) + \sum_{j=1}^m \mu_j \nabla \tilde{g}_{j0}^U(x^0) = 0$;
(KT-2) $\mu_j \cdot \tilde{g}_{j0}^U(x^0) = 0$ for all j = 1,..,m.

But $F(x) = \int_0^1 \tilde{f}_\alpha^L(x)d\alpha + \int_0^1 \tilde{f}_\alpha^U(x)d\alpha$. We obtain the kuhn-Tucker conditions for problem (FOP2) as follows

(FKT-1) $\int_0^1 \nabla \tilde{f}_\alpha^L(x^0)\, d\alpha + \int_0^1 \nabla \tilde{f}_\alpha^U(x^0)\, d\alpha + \sum_{j=1}^m \mu_j \nabla \tilde{g}_{j0}^U(x^0) = 0$;
(FKT-2) $\mu_j \cdot \tilde{g}_{j0}^U(x^0) = 0$ for all j = 1,..,m.

Sufficient. We are going to prove this part by contradiction. Suppose that x^0 not a nondominated solution. Then there exists a $x^1(\neq x^0) \in X$ such that $\tilde{f}(x^1) \prec \tilde{f}(x^0)$. Therefore, we have

$$\tilde{f}_\alpha^L(x^1) + \tilde{f}_\alpha^U(x^1) < \tilde{f}_\alpha^L(x^0) + \tilde{f}_\alpha^U(x^0)$$

for all $\alpha \in [0,1]$. From (5.1), we obtain

$$F(x^1) < F(x^0) \tag{5.2}$$

Since F is convex and continuously differentiable function. Furthermore, $x^0 \in X = \{x \in T \subset \mathbb{R}^n : \tilde{g}_{j0}^U(x) \leq 0, j = 1,..,m\}$, by conditions (FKT-1) and (FKT-2) of this theorem, we obtain the following new conditions:

(KT-1) $\nabla F(x^0) + \sum_{j=1}^m \mu_j \nabla \tilde{g}_{j0}^U(x^0) = 0$;
(KT-2) $\mu_j \cdot \tilde{g}_{j0}^U(x^0) = 0$ for all j = 1,..,m.

Using Theorem 5, we say that x^0 is an optimal solution of real-objective function F with real constraints $\tilde{g}_{j0}^U(x) \leq 0$, for $j = 1,..,m$. i.e., $F(x^0) \leq F(x^1)$, which contradicts to (5.2). Hence the proof. □

We consider here first fuzzy optimization problem having fuzzy-valued objective function and real constraints.

Example 3

$$\begin{aligned} Minimize \quad & \tilde{f}(x_1,x_2) = (\tilde{a} \odot x_1^2) \oplus (\tilde{b} \odot x_2^2) \\ subject\ to\ & g(x_1,x_2) = (x_1-2)^2 + (x_2-2)^2 \leq 1, \end{aligned}$$

where $\tilde{a} = (1,2,3)$ and $\tilde{b} = (0,1,2)$ are triangular fuzzy numbers defined on $\mathbb{R}$ as

$$\tilde{a}(r) = \begin{cases} (r-1), & if\ 1 \leq r \leq 2, \\ (3-r), & if\ 2 < r \leq 3, \\ 0 & otherwise \end{cases} \quad \tilde{b}(r) = \begin{cases} r, & if\ 0 \leq r \leq 1, \\ 2-r, & if\ 1 < r \leq 2, \\ 0 & otherwise \end{cases}$$

Using definition 2, we obtain

$\tilde{f}_\alpha^L(x_1,x_2) = (1+\alpha)x_1^2 + \alpha x_2^2$ and $\tilde{f}_\alpha^U(x_1,x_2) = (3-\alpha)x_1^2 + (2-\alpha)x_2^2$
for $\alpha \in [0,1]$.

We obtain

$\nabla \tilde{f}_\alpha^L(x_1,x_2) = \begin{pmatrix} 2x_1(\alpha+1) \\ 2x_2\alpha \end{pmatrix}$, $\nabla \tilde{f}_\alpha^U(x_1,x_2) = \begin{pmatrix} 2x_1(3-\alpha) \\ 2x_2(2-\alpha) \end{pmatrix}$ and
$\nabla g(x_1,x_2) = \begin{pmatrix} 2(x_1-2) \\ 2(x_2-2) \end{pmatrix}$

Therefore, we have

$$\int_0^1 \nabla \tilde{f}_\alpha^L(x_1,x_2)\, d\alpha = \begin{pmatrix} 3x_1 \\ x_2 \end{pmatrix}, \int_0^1 \nabla \tilde{f}_\alpha^U(x_1,x_2)\, d\alpha = \begin{pmatrix} 5x_1 \\ 3x_2 \end{pmatrix}.$$

From Theorem 6, we have the following Kuhn-Tucker conditions

(FKT-1) $8x_1 + 2\mu(x_1 - 2) = 0,\ 4x_2 + 2\mu(x_2 - 2) = 0,$
(FKT-2) $\mu((x_1 - 2)^2 + (x_2 - 2)^2 - 1) = 0.$

Solving these equations, we get the solution $(x_1, x_2) = (6/5, 3/2)$ and $\mu = 6$. By Theorem 6, we say that $(x_1^*, x_2^*) = (6/5, 3/2)$ is nondominated solution for given problem . Also the minimum value of objective function is $\tilde{f}_{min} = (1.44, 5.13, 8.82)$ and we can find its defuzzified value 5.13 by using center of area method (ref. [7]).

Now we solve the same fuzzy optimization problem having fuzzy-valued objective function with fuzzy constraints.

Example 4

$$\begin{aligned} Minimize \quad & \tilde{f}(x_1, x_2) = (\tilde{a} \odot x_1^2) \oplus (\tilde{b} \odot x_2^2) \\ subject\ to\ & \tilde{g}(x_1, x_2) = (\tilde{b} \odot (x_1 - 2)^2) \oplus (\tilde{b} \odot (x_2 - 2)^2) \preceq \tilde{c}, \end{aligned}$$

where $\tilde{a} = (1, 2, 3)$, $\tilde{b} = (0, 1, 2)$ and $\tilde{c} = (0, 2, 4)$ are triangular fuzzy numbers defined on $\mathbb{R}$ as

$$\tilde{a}(r) = \begin{cases} (r-1), & if\ 1 \le r \le 2, \\ (3-r), & if\ 2 < r \le 3, \\ 0 & otherwise \end{cases} \quad \tilde{b}(r) = \begin{cases} r, & if\ 0 \le r \le 1, \\ 2-r, & if\ 1 < r \le 2, \\ 0 & otherwise \end{cases}$$

$$\tilde{c}(r) = \begin{cases} r/2, & if\ 0 \le r \le 2, \\ (4-r)/2, & if\ 2 < r \le 4, \\ 0 & otherwise \end{cases}$$

Using definition 2, we obtain

$\tilde{f}_\alpha^L(x_1, x_2) = (1 + \alpha)x_1^2 + \alpha x_2^2$ and $\tilde{f}_\alpha^U(x_1, x_2) = (3 - \alpha)x_1^2 + (2 - \alpha)x_2^2$ for $\alpha \in [0, 1]$.

Moreover, $\tilde{g}_\alpha^U(x_1, x_2) = (2 - \alpha)(x_1 - 2)^2 + (2 - \alpha)(x_2 - 2)^2 \le (4 - 2\alpha)$ for $\alpha \in [0, 1]$.

Therefore, $\tilde{g}_0^U(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 2)^2 \le 2$.
Now we obtain

$\nabla \tilde{f}_\alpha^L(x_1, x_2) = \begin{pmatrix} 2x_1(\alpha + 1) \\ 2x_2\alpha \end{pmatrix}$, $\nabla \tilde{f}_\alpha^U(x_1, x_2) = \begin{pmatrix} 2x_1(3 - \alpha) \\ 2x_2(2 - \alpha) \end{pmatrix}$ and
$\nabla g(x_1, x_2) = \begin{pmatrix} 2(x_1 - 2) \\ 2(x_2 - 2) \end{pmatrix}$

Therefore, we have

$$\int_0^1 \nabla \tilde{f}_\alpha^L(x_1, x_2)\, d\alpha = \begin{pmatrix} 3x_1 \\ x_2 \end{pmatrix}, \int_0^1 \nabla \tilde{f}_\alpha^U(x_1, x_2)\, d\alpha = \begin{pmatrix} 5x_1 \\ 3x_2 \end{pmatrix}.$$

From Theorem 6, we have the following Kuhn-Tucker conditions

(FKT-1) $8x_1 + 2\mu(x_1 - 2) = 0,\ 4x_2 + 2\mu(x_2 - 2) = 0,$
(FKT-2) $\mu((x_1 - 2)^2 + (x_2 - 2)^2 - 2) = 0.$

Solving these equations, we get the solution $(x_1, x_2) = ((-6+2\sqrt{41})/(1+\sqrt{41}), (-6+2\sqrt{41})/(-1+\sqrt{41}))$ and $\mu = -3+\sqrt{41}$. By Theorem 6, we say that $(x_1^*, x_2^*) = ((-6+2\sqrt{41})/(1+\sqrt{41}), (-6+2\sqrt{41})/(-1+\sqrt{41}))$ is nondominated solution for given problem . Also the minimum value of objective function is $\tilde{f}_{min} = (0.8453, 3.2773, 5.7094)$ and we can find its defuzzified value 3.2773 by using center of area method.

Remark 2 By comparing the defuzzified value of minimum objective functions in example 3 and 4, we observe that there is significant effect on minimum value of the fuzzy-valued objective function if consider fuzzy optimization problem with fuzzy constraints. Moreover, if we consider the fuzzy optimization problem having fuzzy constraints then we can not apply Theorem 6.2 from [10] to find the nondominated solution. In that case, our result is quite useful to get the solution.

6 Conclusion

Using partial order relation on fuzzy number space, the necessary and sufficient Kuhn-Tucker like optimality condtions for nonlinear fuzzy optimization problem have been derived in this paper. We have used hukuhara differentiability and convexity of fuzzy-valued function for proving the same. We have also provided an example to illustrate the possible applications in this subject.

References

1. Bellman,R.E.and Zadeh,L.A., Decision making in a fuzzy environment, Management Science, 17, 1970,pp. 141-164.
2. Brian S. Thomson, Judith B. Bruckner, Andrew M. Bruckner, Elementary Real Analysis. Prentice Hall (Pearson) Publisher, 2001.
3. Degang,C. and Liguo,Z., Signed Fuzzy valued Measures and Radon-Nikodym Theorem of Fuzzy valued Measurable Functions, Southest Asian Bulletin of Mathematics, 2002.
4. Diamond P., Kloeden, Meric spaces of fuzzy sets: Theory and Applications ,World Scientific, 1994.
5. Chong,E. K. P. and Zak,S. H., An Introduction to Optimization, A Wiley-Interscience Publication, 2005.
6. George,A. A., Fuzzy Ostrowski Type Inequalities, Computational and Applied mathematics, Vol. 22, 2003, pp. 279-292.
7. George,J.K., Bo Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall of India, 1995.
8. Hsien-Chung Wu, Duality Theory in Fuzzy Optimization Problems, Fuzzy Optimization and Decision Making,3, 2004, pp. 345-365.
9. Hsien-Chung Wu, An (α, β)-Optimal Solution Concept in Fuzzy Optimization Problems, Optimization,Vol. 53,April 2004, pp.203-221.
10. Hsien-Chung Wu, The Optimality Conditions for Optimization Problems with Fuzzy-valued Objective Functions, Optimization 57, 2008, pp.473-489.
11. Puri M. L., Ralescu D. A., Differentials of fuzzy functions, J. of Math. Analysis and App., 91, 1983, pp.552-558.
12. Rangarajan K. Sundaram, A First Course in Optimization Theory ,Cambridge University Press, 1996.
13. Rudin, W. ,Principles of Mathematical Analysis (3rd Ed.), New York: McGraw-Hill Book Company, 1976.
14. Saito, S. and Ishii, H., L-Fuzzy Optimization Problems by Parametric Representation,IEEE,2001.
15. Song S., Wu C., Existence and uniqueness of solutions to cauchy problem of fuzzy differential equations, Fuzzy Sets and Systems, 110, 2000, pp. 55-67.
16. Zimmermann, H.J., Fuzzy Set Theory and Its Applications(2nd Ed.),Kluwer-Nijhoff, Hinghum, 1991.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 1, pp. 17 - 32, 2011

SUBCLASSES OF P-VALENT STARLIKE FUNCTIONS DEFINED BY USING CERTAIN FRACTIONAL DERIVATIVE OPERATOR

SOMIA MUFTAH AMSHERI

Department of Mathematics, SCIM, University of Bradford, Bradford, BD7 1DP, U.K.
somia_amsheri@yahoo.com

VALENTINA ZHARKOVA

Department of Mathematics, SCIM, University of Bradford, Bradford, BD7 1DP, U.K.
v.v.zharkova@Bradford.ac.uk

In this paper we introduce two subclasses of p-valent starlike functions in the open unit disk by using fractional derivative operator. We obtain coefficient inequalities and distortion theorems for functions belonging to these subclasses. Further results include distortion theorems (involving the generalized fractional derivative operator). The radii of convexity for functions belonging to these subclasses are also studied.

Keywords: p-valent functions; starlike functions; convex functions; fractional derivative operators.

1. Introduction

Fractional calculus operators have recently found interesting applications in the theory of analytic functions. The classical definition of fractional calculus and its other generalizations have fruitfully been applied in obtaining, the characterization properties, coefficient estimates and distortion inequalities for various subclasses of analytic functions. For numerous references on the subject, one may refer to the works by Srivastava and Owa (1989) and Srivastava and Owa (1992). See also Owa (1978); Srivastava and Owa (1987); Raina and Srivastava (1996) and Raina and Nahar (2002).

Let $A(p)$ denote the class of functions defined by

$$f(z) = z^p + \sum_{n=1}^{\infty} a_{p+n}\, z^{p+n} \qquad (p \in \mathbb{N}) \tag{1.1}$$

which are analytic and p-valent in the unit disk $\mathcal{U} = \{z: |z| < 1\}$. Also denote $T(p)$, the subclass of $A(p)$ consisting of functions of the form

$$f(z) = z^p - \sum_{n=1}^{\infty} a_{p+n}\, z^{p+n} \qquad \left(a_{p+n} \geq 0, \quad p \in \mathbb{N}\right) \tag{1.2}$$

Two interesting subclasses $T^*(\alpha,\beta,\gamma)$ and $C(\alpha,\beta,\gamma)$ of univalent starlike functions with negative coefficients in the open unit disk $\mathcal{U}$ were introduced by Srivastava and Owa (1991a). In fact, these classes become the subclasses of the class $K(\alpha,\beta)$ which was introduced by Gupta (1984) when the function $f(z)$ is univalent with negative coefficients. Using the results of Srivastava and Owa (1991a), Srivastava and Owa (1991b) have obtained several distortion theorems involving fractional derivatives and fractional integrals of functions belonging to the classes $T^*(\alpha,\beta,\gamma)$ and $C(\alpha,\beta,\gamma)$.

Recently, Aouf and Hossen (2006) have generalized the results of Srivastava and Owa (1991a) to the case of certain p-valent functions with negative coefficients.

With these points in view, by making use of a certain fractional derivative operator, we introduce new subclasses $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ (defined below) of p-valent starlike functions with negative coefficients.

This paper is organized as follows: Section 2 gives preliminary details and definitions of p-valent starlike functions, p-valent convex functions and fractional derivative operators. In Section 3 we describe coefficient inequalities for the functions belonging to the subclasses $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Section 4 considers the distortion properties. Its further distortion properties (involving the generalized fractional derivative operator) are discussed in section 5. Finally, in section 6 we determine the radii of convexity for functions belonging to these subclasses of p-valent starlike functions.

2. Preliminaries And Definitions

A function $f(z) \in A(p)$ is called p-valent starlike of order α if $f(z)$ satisfies the conditions

$$\mathrm{Re}\left\{\frac{zf'(z)}{f(z)}\right\} > \alpha \tag{2.1}$$

and

$$\int_0^{2\pi} \mathrm{Re}\left\{\frac{zf'(z)}{f(z)}\right\} d\theta = 2p\pi \tag{2.2}$$

for $0 \le \alpha < p,\ p \in \mathbb{N}$, and $z \in \mathcal{U}$. We denote by $S^*(p,\alpha)$ the class of all p-valent starlike functions of order α. Also a function $f(z) \in A(p)$ is called p-valent convex of order α if $f(z)$ satisfies the following conditions:

$$\mathrm{Re}\left\{1+\frac{zf''(z)}{f'(z)}\right\} > \alpha \tag{2.3}$$

and

$$\int_0^{2\pi} \mathrm{Re}\left\{1+\frac{zf''(z)}{f'(z)}\right\} d\theta = 2p\pi \tag{2.4}$$

for $0 \leq \alpha < p,\ p \in \mathbb{N}$, and $z \in \mathcal{U}$. We denote by $K(p,\alpha)$ the class of all p-valent convex functions of order α. We note that

$$f(z) \in K(p,\alpha) \Leftrightarrow \frac{zf'(z)}{p} \in S^*(p,\alpha) \tag{2.5}$$

for $0 \leq \alpha < p$.

The class $S^*(p,\alpha)$ was introduced by Patil and Thakare (1983), and the class $K(p,\alpha)$ was introduced by Owa (1985). Also, we denote by $T^*(p,\alpha)$ and $C(p,\alpha)$ the classes obtained by taking intersections, respectively, of the classes $S^*(p,\alpha)$ and $K(p,\alpha)$ with $T(p)$; that is

$$T^*(p,\alpha) = S^*(p,\alpha) \cap T(p)$$

and

$$C(p,\alpha) = K(p,\alpha) \cap T(p)$$

The classes $T^*(p,\alpha)$ and $C(p,\alpha)$ were introduced by Owa (1985). In particular, the classes $T^*(1,\alpha) = T^*(\alpha)$ and $C(1,\alpha) = C(\alpha)$ when $p = 1$, were studied by Silverman (1975).

Let ${}_2F_1(a,b;c;z)$ be the Gauss hypergeometric function defined, for $z \in \mathcal{U}$ by; see Srivastava and Karlsson (1985).

$${}_2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n\,(b)_{n\ z^n}}{(c)_n\ n!} \tag{2.6}$$

where $(\lambda)_n$ is the Pochhammer symbol defined, in terms of the Gamma function, by

$$(\lambda)_n = \frac{\Gamma(\lambda+n)}{\Gamma(\lambda)} = \begin{cases} 1 & , \quad n = 0 \\ \lambda(\lambda+1)(\lambda+2)\ldots.(\lambda+n-1) & , \quad n \in \mathbb{N} \\ (\lambda \neq 0, -1, -2, \ldots) & \end{cases} \tag{2.7}$$

Now we recall the following definitions of fractional derivative operators, adopted for working in classes of analytic functions in complex plane which were used by Owa (1978); Srivastava and Owa (1987); Srivastava and Owa (1991b); Raina and Nahar (2002); Raina and Srivastava (1996); see also, Srivastava and Owa (1989) and Srivastava and Owa (1992) as follows.

Definition 2.1. *The fractional derivative of order λ is defined, for a function $f(z)$, by*

$$D_z^{\lambda} f(z) = \frac{1}{\Gamma(1-\lambda)} \frac{d}{dz} \int_0^z \frac{f(\xi)}{(z-\xi)^{\lambda}}\, d\xi \tag{2.8}$$

where $0 \leq \lambda < 1$ *is analytic function in a simply- connected region of the z-plane containing the origin, and the multiplicity of* $(z-\xi)^{-\lambda}$ *is removed by requiring log* $(z-\xi)$ *to be real when* $z-\xi > 0$.

Definition 2.2. *Let* $0 \leq \lambda < 1$, *and* $\mu, \eta \in \mathbb{R}$. *Then, in terms of the familiar Gauss's hypergeometric function* ${}_2F_1$, *the generalized fractional derivative operator* $J_{0,Z}^{\lambda,\mu,\eta}$ *is*

$$J_{0,Z}^{\lambda,\mu,\eta} f(z) = \frac{d}{dz}\left(\frac{z^{\lambda-\mu}}{\Gamma(1-\lambda)}\int_0^z (z-\xi)^{-\lambda} f(\xi)\, {}_2F_1\left(\mu-\lambda, 1-\eta; 1-\lambda; 1-\frac{\xi}{z}\right) d\xi\right) \quad (2.9)$$

where f(z) is analytic function in a simply- connected region of the z-plane containing the origin with the order $f(z) = O(|z|^{\varepsilon})$, $z \to 0$, *where* $\varepsilon > max\{0, \mu-\eta\} - 1$, *and the multiplicity of* $(z-\xi)^{-\lambda}$ *is removed by requiring log* $(z-\xi)$ *to be real when* $z-\xi > 0$.

Notice that

$$J_{0,Z}^{\lambda,\lambda,\eta} f(z) = D_z^{\lambda} f(z), \qquad 0 \leq \lambda < 1 \quad (2.10)$$

Srivastava and Owa (1987) have established various distortion theorems for the fractional calculus of functions $f(z)$ belonging to the classes $T^*(p,\alpha)$ and $C(p,\alpha)$.

We now define the following classes of p-valent starlike functions based on fractional derivative operator.

Definition 2.3. *The function* $f(z) \in T(p)$ *is said to be in the class* $T_{\lambda,\mu,\eta}^*(p,\alpha,\beta,\gamma)$ *if*

$$\left|\frac{\dfrac{z\left(\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right)'}{g(z)} - p}{\dfrac{z\left(\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right)'}{g(z)} + p - 2\beta}\right| < \gamma \qquad (z \in \mathcal{U}) \quad (2.11)$$

$(\lambda \geq 0, \mu < p+1\,;\ \eta > \max(\lambda,\mu) - p - 1\,;\ 0 \leq \alpha < p;\ 0 \leq \beta < p;\ 0 < \gamma \leq 1;\ p \in \mathbb{N})$

For the function

$$g(z) = z^p - \sum_{n=1}^{\infty} b_{p+n}\, z^{p+n} \qquad \left(b_{p+n} \geq 0\,, p \in \mathbb{N}\right) \quad (2.12)$$

belonging to $T^*(p,\alpha)$. *Dented by* $M_{0,z}^{\lambda,\mu,\eta} f(z)$ *the modification of the fractional derivative operator which is defined in terms of* $J_{0,Z}^{\lambda,\mu,\eta}$ *as follows:*

$$\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z) = \phi_p(\lambda,\mu,\eta)\, z^{\mu}\, J_{0,Z}^{\lambda,\mu,\eta} f(z) \quad (2.13)$$

with

$$\phi_p(\lambda,\mu,\eta) = \frac{\Gamma(1-\mu+p)\Gamma(1+\eta-\lambda+p)}{\Gamma(1+p)\Gamma(1+\eta-\mu+p)} \quad (2.14)$$

Further, if $f(z) \in T(p)$ *satisfies the condition (2.11) for* $g(z) \in C(p,\alpha)$, *we say that* $f(z) \in C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$.

The above-defined classes $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ are of special interest and they contain many well-known classes of analytic functions. In particular, in view of (2.13), we find that

$$\mathrm{M}_{0,z}^{0,0,\eta} f(z) = f(z) \tag{2.15}$$

Thus, for $\lambda = \mu = 0$, we have

$$T^*_{0,0,\eta}(p,\alpha,\beta,\gamma) = T^*(p,\alpha,\beta,\gamma)$$

and

$$C_{0,0,\eta}(p,\alpha,\beta,\gamma) = C(p,\alpha,\beta,\gamma)$$

where $T^*(p,\alpha,\beta,\gamma)$ and $C(p,\alpha,\beta,\gamma)$ are precisely the subclasses of p-valent starlike functions which were studied by Aouf and Hossen (2006).

Furthermore, for $\lambda = \mu = 0$ and $p = 1$, we obtain

$$T^*_{0,0,\eta}(1,\alpha,\beta,\gamma) = T^*(\alpha,\beta,\gamma)$$

and

$$C_{0,0,\eta}(1,\alpha,\beta,\gamma) = C(\alpha,\beta,\gamma)$$

where $T^*(\alpha,\beta,\gamma)$ and $C(\alpha,\beta,\gamma)$ are the subclasses of starlike functions which were studied by Srivastava and Owa (1991a) as well as by Srivastava and Owa (1991b).

In order to prove our results we shall need the following lemmas for the classes $T^*(p,\alpha)$ and $C(p,\alpha)$ due to Owa (1985):

Lemma 2.4. *Let the function* $g(z)$ *defined by (2.12). Then* $g(z)$ *is in the class* $T^*(p,\alpha)$ *if and only if*

$$\sum_{n=1}^{\infty} (p+n-\alpha) b_{p+n} \leq (p-\alpha) \tag{2.16}$$

Lemma 2.5. *Let the function* $g(z)$ *defined by (2.12). Then* $g(z)$ *is in the class* $C(p,\alpha)$ *if and only if*

$$\sum_{n=1}^{\infty} (p+n)(p+n-\alpha) b_{p+n} \leq p(p-\alpha) \tag{2.17}$$

We mention to the following known result which shall be used in the sequel (Raina and Nahar (2002); see also Raina and Srivastava (1996)).

Lemma 2.6. *Let* $\lambda, \mu, \eta \in \mathbb{R}$, *such that* $\lambda \geq 0$ *and* $k > max\{0, \mu - \eta\} - 1$ *then*

$$J_{0,Z}^{\lambda,\mu,\eta} z^k = \frac{\Gamma(k+1)\Gamma(k-\mu+\eta+1)}{\Gamma(k-\mu+1)\Gamma(k-\lambda+\eta+1)} z^{k-\mu} \tag{2.18}$$

3. Coefficient Inequalities

Theorem 3.1. *Let the function $f(z)$ be defined by (1.2). If $f(z)$ belongs to the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$, then*

$$\sum_{n=1}^{\infty} \delta_n(\lambda,\mu,\eta,p)(p+n)\,(1+\gamma)a_{p+n} - \frac{(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{p+n-\alpha} \le 2\gamma(p-\beta) \tag{3.1}$$

where

$$\delta_n(\lambda,\mu,\eta,p) = \frac{\phi_p(\lambda,\mu,\eta)}{\phi_{p+n}(\lambda,\mu,\eta)} = \frac{(1+p)_n(1+\eta-\mu+p)_n}{(1-\mu+p)_n(1+\eta-\lambda+p)_n} \tag{3.2}$$

and $\phi_p(\lambda,\mu,\eta)$ is given by (2.14).

Proof. Applying Lemma 2.6, we have from (1.2) and (2.13) that

$$\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z) = z^p - \sum_{n=1}^{\infty} \frac{(1+p)_n(1+\eta-\mu+p)_n}{(1-\mu+p)_n(1+\eta-\lambda+p)_n}\, a_{p+n}\, z^{p+n}$$

Since $f(z) \in T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$, there exist a function $g(z)$ belonging to the class $T^*(p,\alpha)$ such that

$$\left| \frac{z\left(\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right)' - pg(z)}{z\left(\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right)' + (p-2\beta)g(z)} \right| < \gamma, \quad z \in \mathcal{U} \tag{3.3}$$

It follows from (3.3) that

$$\mathrm{Re}\left\{ \frac{\sum_{n=1}^{\infty}[\delta_n(\lambda,\mu,\eta,p)(p+n)\,a_{p+n} - p\,b_{p+n}]\; z^n}{2(p-\beta) - \sum_{n=1}^{\infty}[\delta_n(\lambda,\mu,\eta,p)(p+n)a_{p+n} + (p-2\beta)b_{p+n}]z^n} \right\} < \gamma \tag{3.4}$$

Choosing values of z on the real axis so that $\frac{z\left(\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right)'}{g(z)}$ is real, and letting $z \to 1^-$ through real axis, we have

$$\sum_{n=1}^{\infty}[\delta_n(\lambda,\mu,\eta,p)(p+n)\,a_{p+n} - p\,b_{p+n}] \le$$

$$\gamma\left\{2(p-\beta) - \sum_{n=1}^{\infty}[\delta_n(\lambda,\mu,\eta,p)(p+n)a_{p+n} + (p-2\beta)b_{p+n}]\right\}$$

or, equivalently,

$$\sum_{n=1}^{\infty}\{\delta_n(\lambda,\mu,\eta,p)(p+n)\,(1+\gamma)a_{p+n} - [p(1-\gamma)+2\gamma\beta]\,b_{p+n}\} \le 2\gamma(p-\beta) \quad (3.5)$$

Note that, by using Lemma 2.4, $g(z) \in T^*(p,\alpha)$ implies

$$b_{p+n} \le \frac{p-\alpha}{p+n-\alpha} \quad (3.6)$$

Making use of (3.6) in (3.5), we complete the proof of Theorem 3.1. □

Corollary 3.2. *Let the function $f(z)$ be defined by (1.2) be in the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then*

$$a_{p+n} \le \frac{2\gamma(p-\beta)(p+n-\alpha)+(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{\delta_n(\lambda,\mu,\eta,p)(p+n)(1+\gamma)(p+n-\alpha)} \quad (3.7)$$

where $\delta_n(\lambda,\mu,\eta,p)$ is given by (3.2). The result (3.7) is sharp for a function of the form:

$$f(z) = z^p - \frac{2\gamma(p-\beta)(p+n-\alpha)+(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{\delta_n(\lambda,\mu,\eta,p)(p+n)(1+\gamma)(p+n-\alpha)} z^{p+n} \quad (3.8)$$

with respect to

$$g(z) = z^p - \frac{p-\alpha}{p+n-\alpha} z^{p+n}, \qquad n \ge 1 \quad (3.9)$$

Remark 1. Letting $p = 1$, $\lambda = \mu = 0$, and $\alpha = 0$ in Corollary 3.2, we obtain a result was proved by [Gupta (1984), Theorem 3].

In a similar manner, Lemma 2.5 can be used to prove the following theorem:

Theorem 3.3. *Let the function $f(z)$ be defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then*

$$\sum_{n=1}^{\infty} \delta_n(\lambda,\mu,\eta,p)(p+n)\,(1+\gamma)a_{p+n} - \frac{p(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{(p+n)(p+n-\alpha)} \le 2\gamma(p-\beta) \quad (3.10)$$

where $\delta_n(\lambda,\mu,\eta,p)$ is given by (3.2).

Corollary 3.4. *Let the function $f(z)$ be defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then*

$$a_{p+n} \le \frac{2\gamma(p-\beta)(p+n)(p+n-\alpha)+p(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{\delta_n(\lambda,\mu,\eta,p)(p+n)^2(1+\gamma)(p+n-\alpha)} \quad (3.11)$$

where $\delta_n(\lambda,\mu,\eta,p)$ is given by (3.2). The result (3.11) is sharp for a function of the form:

$$f(z) = z^p - \frac{2\gamma(p-\beta)(p+n)(p+n-\alpha) + p(p-\alpha)[p(1-\gamma)+2\gamma\beta]}{\delta_n(\lambda,\mu,\eta,p)(p+n)^2(1+\gamma)(p+n-\alpha)} z^{p+n} \qquad (3.12)$$

with respect to

$$g(z) = z^p - \frac{p(p-\alpha)}{(p+n)(p+n-\alpha)} z^{p+n}, \qquad n \geq 1 \qquad (3.13)$$

4. Distortion Properties

Next, we state and prove results concerning distortion properties of $f(z)$ belonging to the classes $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$.

Theorem 4.1. *Let* $\lambda, \mu, \eta \in \mathbb{R}$ *such that*

$$\lambda \geq 0, \mu < p+1\,, \ \ \eta \geq \lambda\left(1 - \frac{p+2}{\mu}\right) \ and \ \ p \in \mathbb{N} \qquad (4.1)$$

Also, let $f(z)$ *defined by (1.2) be in the class* $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. *Then*

$$|f(z)| \geq |z|^p - A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)\,|z|^{p+1}, \qquad (4.2)$$

$$|f(z)| \leq |z|^p + A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^{p+1}, \qquad (4.3)$$

$$|f'(z)| \geq p|z|^{p-1} - (\mathrm{p}+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^p, \qquad (4.4)$$

and

$$|f'(z)| \leq p|z|^{p-1} + (\mathrm{p}+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^p, \qquad (4.5)$$

for $z \in \mathcal{U}$, *provided that* $0 \leq \alpha < p, 0 \leq \beta < p$ *and* $0 < \gamma \leq 1$, *where*

$$A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) = \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)\}}{(1+p+\eta-\mu)(p+1)^2(1+\gamma)(p+1-\alpha)} \qquad (4.6)$$

The estimates for $|f(z)|$ *and* $|f'(z)|$ *are sharp.*

Proof. We observe that the function $\delta_n(\lambda,\mu,\eta,p)$ defined by (3.2) satisfy the inequality $\delta_n(\lambda,\mu,\eta,p) \leq \delta_{n+1}(\lambda,\mu,\eta,p)$, $\forall n \in \mathbb{N}$, provided that $\eta \geq \lambda\left(1 - \frac{p+2}{\mu}\right)$. Thereby, showing that $\delta_n(\lambda,\mu,\eta,p)$ is non-decreasing. Thus under conditions stated in (4.1), we have

$$0 < \frac{(1+p)(1+p+\eta-\mu)}{(1+p-\mu)(1+p+\eta-\lambda)} = \delta_1(\lambda,\mu,\eta,p) \leq \delta_n(\lambda,\mu,\eta,p), \qquad \forall n \in \mathbb{N} \qquad (4.7)$$

For $f(z) \in T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$, (3.5) implies

$$\delta_1(\lambda,\mu,\eta,p)(p+1)(1+\gamma)\sum_{n=1}^{\infty} a_{p+n} - [p(1-\gamma)+2\gamma\beta]\sum_{n=1}^{\infty} b_{p+n} \leq 2\gamma(p-\beta) \qquad (4.8)$$

For $g(z) \in T^*(p,\alpha)$, Lemma 2.4 yields

$$\sum_{n=1}^{\infty} b_{p+n} \leq \frac{p-\alpha}{p+1-\alpha} \tag{4.9}$$

So that (4.8) reduces to

$$\sum_{n=1}^{\infty} a_{p+n} \leq \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)\}}{(1+p+\eta-\mu)(p+1)^2(1+\gamma)(p+1-\alpha)}$$

$$= A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) \tag{4.10}$$

Consequently,

$$|f(z)| \geq |z|^p - |z|^{p+1}\sum_{n=1}^{\infty} a_{p+n} \tag{4.11}$$

and

$$|f(z)| \leq |z|^p + |z|^{p+1}\sum_{n=1}^{\infty} a_{p+\mathrm{n}} \tag{4.12}$$

On using (4.11), (4.12) and (4.10), we easily arrive at the desired results (4.2) and (4.3).

Furthermore, we note from (3.5) that

$$\delta_1(\lambda,\mu,\eta,p)(1+\gamma)\sum_{n=1}^{\infty}(p+n)a_{p+n} - [p(1-\gamma)+2\gamma\beta]\sum_{n=1}^{\infty} b_{p+n} \leq 2\gamma(p-\beta) \tag{4.13}$$

which in view of (4.9), becomes

$$\sum_{n=1}^{\infty}(p+n)a_{p+n} \leq \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)\}}{(1+p+\eta-\mu)(p+1)(1+\gamma)(p+1-\alpha)}$$

$$= (p+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) \tag{4.14}$$

Thus, we have

$$|f'(z)| \geq p|z|^{p-1} - |z|^p\sum_{n=1}^{\infty}(p+n)a_{p+n} \tag{4.15}$$

and

$$|f'(z)| \leq p|z|^{p-1} + |z|^p\sum_{n=1}^{\infty}(p+n)\,a_{p+n} \tag{4.16}$$

On using (4.15), (4.16) and (4.14), we arrive at the desired results (4.4) and (4.5).

Finally, we can prove that the estimates for $|f(z)|$ and $|f'(z)|$ are sharp by taking the function

$$f(z) = z^p - \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)\}}{(1+p+\eta-\mu)(p+1)^2(1+\gamma)(p+1-\alpha)} z^{p+1} \quad (4.17)$$

with respect to

$$g(z) = z^p - \frac{p-\alpha}{p+1-\alpha} z^{p+1}, \qquad n \geq 1 \quad (4.18)$$

This completes the proof of Theorem 4.1. □

Corollary 4.2. *Let the function $f(z)$ be defined by (1.2) be in the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then the unit disk $\mathcal{U}$ is mapped onto a domain that contains the disk $|w| < r_1$, where*

$$r_1 = 1 - \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)\}}{(1+p+\eta-\mu)(p+1)^2(1+\gamma)(p+1-\alpha)} \quad (4.19)$$

The result is sharp with the extremal function defined by (4.17).

Remark 2. Letting $p = 1$, $\lambda = \mu = 0$, and $\alpha = 0$ in Theorem 4.1, we obtain a result was proved by [Gupta (1984), Theorem 4].

Theorem 4.3. *Under the conditions stated in (4.1), let the function $f(z)$ defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then*

$$|f(z)| \geq |z|^p - B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^{p+1}, \quad (4.20)$$

$$|f(z)| \leq |z|^p + B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^{p+1}, \quad (4.21)$$

$$|f'(z)| \geq p|z|^{p-1} - (p+1)B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^p, \quad (4.22)$$

and

$$|f'(z)| \leq p|z|^{p-1} + (p+1)B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)|z|^p, \quad (4.23)$$

for $z \in \mathcal{U}$, provided that $0 \leq \alpha < p, 0 \leq \beta < p$ and $0 < \gamma \leq 1$, where

$$B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) = \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)(p+1)(p+1-\alpha)+p(p-\alpha)[p(1-\gamma)+2\gamma\beta]\}}{(1+p+\eta-\mu)(p+1)^3(1+\gamma)(p+1-\alpha)} \quad (4.24)$$

The estimates for $|f(z)|$ and $|f'(z)|$ are sharp.

Proof. By using Lemma 2.5, we have

$$\sum_{n=1}^{\infty} b_{p+n} \leq \frac{p(p-\alpha)}{(p+1)(p+1-\alpha)} \quad (4.25)$$

since $g(z) \in C(p,\alpha)$, the assertions (4.20), (4.21), (4.22) and (4.23) of Theorem 4.3 follow if we apply (4.25) to (3.5).
The estimates for $|f(z)|$ and $|f'(z)|$ are attained by the function

$$f(z) = z^p - \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)(p+1)(p+1-\alpha)+p(p-\alpha)[p(1-\gamma)+2\gamma\beta]\}}{(1+p+\eta-\mu)(p+1)^3(1+\gamma)(p+1-\alpha)} z^{p+1} \quad (4.26)$$

with respect to

$$g(z) = z^p - \frac{p(p-\alpha)}{(p+1)(p+1-\alpha)} z^{p+1}, \qquad n \geq 1 \tag{4.27}$$

This completes the proof of Theorem 4.3. □

Corollary 4.4. *Let the function $f(z)$ be defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then the unit disk $\mathcal{U}$ is mapped onto a domain that contains the disk $|w| < r_2$, where*

$$r_2 = 1 - \frac{(1+p-\mu)(1+p+\eta-\lambda)\{2\gamma(p-\beta)(\mathrm{p}+1)(p+1-\alpha)+p(p-\alpha)[\mathrm{p}(1-\gamma)+2\gamma\beta]\}}{(1+p+\eta-\mu)(\mathrm{p}+1)^3(1+\gamma)(p+1-\alpha)} \tag{4.28}$$

The result is sharp with the extremal function defined by (4.26).

5. Further Distortion Properties

We next prove two further distortion theorems involving generalized fractional derivative operator $J_{0,Z}^{\lambda,\mu,\eta}$.

Theorem 5.1. *Let $\lambda \geq 0$; $\mu < p+1; \eta > max(\lambda,\mu) - p - 1$, and $p \in \mathbb{N}$. Also let the function $f(z)$ be defined by (1.2) be in the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$.Then*

$$\left|J_{0,Z}^{\lambda,\mu,\eta} f(z)\right| \geq \frac{|z|^{p-\mu}}{\phi_p(\lambda,\mu,\eta)}\left\{1 - \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}|\mathrm{z}|\right\} \tag{5.1}$$

and

$$\left|J_{0,Z}^{\lambda,\mu,\eta} f(z)\right| \leq \frac{|z|^{p-\mu}}{\phi_p(\lambda,\mu,\eta)}\left\{1 + \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}|\mathrm{z}|\right\} \tag{5.2}$$

for $z \in \mathcal{U}$, and $\phi_p(\lambda,\mu,\eta)$ is given by (2.14). The results (5.1) and (5.2) are sharp.

Proof. Consider the function $\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)$ defined by (2.13). With the aid of (4.7) and (4.14) we find that

$$\left|\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right| \geq |z|^p - \delta_1(\lambda,\mu,\eta,p)|z|^{p+1}\sum_{n=1}^{\infty}(p+n)a_{p+n}$$

$$\geq |z|^p - \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}|z|^{p+1} \tag{5.3}$$

and

$$\left|\mathrm{M}_{0,z}^{\lambda,\mu,\eta} f(z)\right| \leq |z|^p + \delta_1(\lambda,\mu,\eta,p)|z|^{p+1} \sum_{n=1}^{\infty} (p+n)a_{p+n}$$

$$\leq |z|^p + \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}|z|^{p+1} \qquad (5.4)$$

which yields the inequality (5.1) and (5.2) of Theorem 5.1.

Finally, by taking the function $f(z)$ defined by

$$J_{0,Z}^{\lambda,\mu,\eta} f(z) = \frac{z^{p-\mu}}{\phi_p(\lambda,\mu,\eta)}\left\{1 - \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}\mathrm{z}\right\} \qquad (5.5)$$

The results (5.1) and (5.2) are easily seen to be sharp. □

Corollary 5.2. *Let the function $f(z)$ defined by (1.2) be in the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then $J_{0,Z}^{\lambda,\mu,\eta} f(z)$ is included in a disk with its centre at the origin and radius r_3 given by*

$$r_3 = \frac{1}{\phi_p(\lambda,\mu,\eta)}\left\{1 + \frac{2\gamma(p-\beta)+p(p-\alpha)(1+\gamma)}{(1+\gamma)(p+1-\alpha)}\right\} \qquad (5.6)$$

Similarly we can establish the following result:

Theorem 5.3. *Let $\lambda \geq 0;\ \mu < p+1; \eta > max(\lambda,\mu) - p - 1,$ and $p \in \mathbb{N}$, and let the function $f(z)$ be defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then*

$$\left|J_{0,Z}^{\lambda,\mu,\eta} f(z)\right| \geq \frac{|z|^{p-\mu}}{\phi_p(\lambda,\mu,\eta)}\left\{1 - \frac{2\gamma(p-\beta)(\mathrm{p}+1)(p+1-\alpha)+p(p-\alpha)[\mathrm{p}(1-\gamma)+2\gamma\beta]}{(1+\mathrm{p})(1+\gamma)(p+1-\alpha)}|\mathrm{z}|\right\} \qquad (5.7)$$

and

$$\left|J_{0,Z}^{\lambda,\mu,\eta} f(z)\right| \leq \frac{|z|^{p-\mu}}{\phi_p(\lambda,\mu,\eta)}\left\{1 + \frac{2\gamma(p-\beta)(\mathrm{p}+1)(p+1-\alpha)+p(p-\alpha)[\mathrm{p}(1-\gamma)+2\gamma\beta]}{(1+\mathrm{p})(1+\gamma)(p+1-\alpha)}|\mathrm{z}|\right\} \qquad (5.8)$$

for $z \in \mathcal{U}$, and $\phi_p(\lambda,\mu,\eta)$ is given by (2.14). The results (5.7) and (5.8) are sharp.

Corollary 5.4. *Let the function $f(z)$ defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then $J_{0,Z}^{\lambda,\mu,\eta} f(z)$ is included in a disk with its centre at the origin and radius r_4 given by*

$$r_4 = \frac{1}{\phi_p(\lambda,\mu,\eta)}\left\{1 + \frac{2\gamma(p-\beta)(\mathrm{p}+1)(p+1-\alpha)+p(p-\alpha)[\mathrm{p}(1-\gamma)+2\beta\gamma]}{(1+\mathrm{p})(1+\gamma)(p+1-\alpha)}\right\} \qquad (5.9)$$

Remark 3. Letting $p = 1, \mu = \lambda$ and using the relationship (2.10) in Theorem 5.1, Corollary 5.2, Theorem 5.3, and Corollary 5.4, we obtain the results which were proved

by [Srivastava and Owa (1991b), Theorem 5, Corollary 3, Theorem 6, and Corollary 4, respectively].

6. Convexity Of Functions In $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ And $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$

In view of Lemma 2.4, we know that the function $f(z)$ defined by (1.2) is p-valent starlike in the unit disk $\mathcal{U}$ if and only if

$$\sum_{n=1}^{\infty}(p+n)a_{p+n} \le p \tag{6.1}$$

for $f(z) \in T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$, we find from (3.5) and (4.9) that

$$\sum_{n=1}^{\infty}(p+n)a_{p+n} \le (p+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) \le p \tag{6.2}$$

where $A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ is defined by (4.6). Furthermore, $f(z) \in C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$, we have

$$\sum_{n=1}^{\infty}(p+n)a_{p+n} \le (p+1)B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma) \le p \tag{6.3}$$

where $B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ is defined by (4.24). Thus we observe that $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ are subclasses of p-valent starlike functions. Naturally, therefore, we are interested in finding the radii of convexity for functions in $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. We first state:

Theorem 6.1. *Let the function $f(z)$ defined by (1.2) be in the class $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then $f(z)$ is p-valent convex in the disk $|z| < r_5$, where*

$$r_5 = \inf_{n\in\mathbb{N}}\left\{\frac{p^2}{(p+1)(p+n)\,A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)}\right\}^{1/n} \tag{6.4}$$

and $A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ is given by (4.6). The result is sharp.

Proof. It suffices to prove

$$\left|1+\frac{zf''(z)}{f'(z)}-p\right| \le p \qquad (|z| < r_5) \tag{6.5}$$

Indeed we have

$$\left|1+\frac{zf''(z)}{f'(z)}-p\right| = \left|\frac{-\sum_{n=1}^{\infty} n(p+n)a_{p+n}z^n}{p-\sum_{n=1}^{\infty}(p+\mathrm{n})a_{p+n}z^n}\right|$$

$$\leq \frac{\sum_{n=1}^{\infty} n(p+n)\, a_{p+n}|z|^n}{p - \sum_{n=1}^{\infty}(p+n)a_{p+n}|z|^n} \tag{6.6}$$

Hence (6.5) is true if

$$\sum_{n=1}^{\infty} n(p+n)\, a_{p+n}|z|^n \leq p^2 - \sum_{n=1}^{\infty} p(p+n)a_{p+n}|z|^n \tag{6.7}$$

that is, if

$$\sum_{n=1}^{\infty} (p+n)^2\, a_{p+n}|z|^n \leq p^2 \tag{6.8}$$

with the aid of (4.14), (6.8) is true if

$$(p+n)|z|^n \leq \frac{p^2}{(p+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)} \tag{6.9}$$

Solving (6.9) for $|z|$, we get

$$|z| \leq \left\{\frac{p^2}{(p+1)(p+n)\, A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)}\right\}^{1/n} \quad (n \geq 1) \tag{6.10}$$

Finally, since $(p+n)^{-1/n}$ is an increasing function for integers $n \geq 1,\ p \in \mathbb{N}$, we have (6.5) for $|z| < r_5$, where r_5 is given by (6.4).

In order to complete the proof of Theorem 6.1, we note that the result is sharp for the function $f(z) \in T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ of the form

$$f(z) = z^p - \frac{(p+1)A_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)}{(p+n)} z^{p+n} \tag{6.11}$$

for some integers $n \geq 1$. □

Similarly we can prove the next theorem.

Theorem 6.2. *Let the function $f(z)$ defined by (1.2) be in the class $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. Then $f(z)$ is p-valent convex in the disk $|z| < r_6$, where*

$$r_6 = \inf_{n\in\mathbb{N}} \left\{\frac{p^2}{(p+1)(p+n)\, B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)}\right\}^{1/n} \tag{6.12}$$

and $B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ is given by (4.24). The result is sharp for the function $f(z) \in C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ of the form

$$f(z) = z^p - \frac{(p+1)B_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)}{(p+n)} z^{p+n} \tag{6.13}$$

for some integers $n \geq 1$.

7. Conclusion

We have studied new classes $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ of p-valent functions with negative coefficients defined by a certain fractional derivative operator in the unit disk $\mathcal{U}$. Many of the known results follow as particular cases from our results; see for example, Aouf and Hossen (2006); Gupta (1984); Srivastava and Owa (1991a) and Srivastava and Owa(1991b). We obtained the sufficient conditions for the function $f(z)$ to be in $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. In addition, we derived a number of distortion theorems of functions belonging to these classes as well as distortion theorems for a certain fractional derivative operator of functions in the classes $T^*_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$ and $C_{\lambda,\mu,\eta}(p,\alpha,\beta,\gamma)$. At the end of this paper we have determined the radii of convexity for functions belonging to these classes

Acknowledgments

The authors thank the referee for some useful suggestions for improvement of the article.

References

Aouf, M. K. ; Hossen, H. M. (2006): Certain subclasses of p-valent starlike functions, Proc. Pakistan Acad. Sci. **43**(2), pp. 99-104.

Gupta, V. P. (1984): convex class of starlike functions, Yokohama Math. J. **32**, pp. 55-59.

Owa, S. (1978): On the distortion theorems. I, Kyungpook Math. J. **18**, pp. 53-59.

Owa, S. (1985): On certain classes of p-valent functions with negative coefficients, Bull. Belg. Math. Soc. Simon stevin, **59**, pp. 385-402.

Raina, R. K. ; Nahar, T. S. (2002): On certain subclasses of p-valent functions defined in terms of certain fractional derivative operators, Glasnik Matematicki, **37**(1), pp. 59-71.

Raina, R. K. ; Srivastava, H. M. (1996): A certain subclass of analytic functions associated with operators of fractional calculus, Comput. Math. Appl. **32**, pp. 13-19.

Partil, D. A. ; Thakare, N. K. (1983): On convex hulls and extreme points of p-valent starlike and convex classes with applications, Bull. Math. Soc. Sci. Math. R. S. Roumanie (N. S.), **27**, pp. 145-160.

Silverman, H. (1975): Univalent functions with negative coefficients, Proc. Amer. Math. Soc. **51**, pp. 109-116.

Srivastava, H. M. ; Karlsson, P. W. (1985): Multiple Gaussian hypergeometric series, Halsted Press (Ellis Horwood Limited, Chichester), Wiley, New York/Chichester/Brisbane/Toronto.

Srivastava, H. M. ; Owa, S. (1987): Some applications of fractional calculus operators to certain classes of analytic and multivalent functions, J. Math. Anal. Appl. **122**, pp. 187-196.

Srivastava, H. M. ; Owa, S. (Eds) (1989): Univalent functions, fractional calculus, and their applications, Halsted Press/Ellis Horwood Limited/John Wiley and Sons, New York/Chichester/Brisbane/Toronto.

Srivastava, H. M. ; Owa, S. (1991a): Certain subclasses of starlike functions - I, J. Math. Anal. Appl. **161**, pp. 405-415.

Srivastava, H. M. ; Owa, S. (1991b): Certain subclasses of starlike functions - II, J. Math. Anal. Appl. **161**, pp. 416-425.

Srivastava, H. M. ; Owa, S. (Eds) (1992): Current topics in analytic function theory, World Scientific Publishing Company, Singapore, New Jersey London and Hong Kong, 1992.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 1, pp. 33 - 48, 2011

ANALYSIS OF THE ROLE OF ENTROPIES AND ASSIGNMENT OF RANKS TO THE FEATURES IN GENRE DISCRIMINATION

HEMLATA PANDE[a] AND H. S. DHAMI[b]

Abstract

Genre or style, an important property of text, and automatic text genre discrimination is becoming important for classification and retrieval purposes as well as for many natural language processing tasks. Various methods with feature cue vectors have been used for genre discrimination, which utilize different statistical measures corresponding to a range of linguistic features. Since all features of a text are not equally useful so for better results and to reduce the dimensionality of used spaces, *it becomes necessary to select those features which have the capacity to discriminate between categories in a better manner. In order to achieve this aim, a technique known as feature selection is commonly applied in text classification. These techniques attempt to select the subset of features in text categorization that are most useful for the categorization task. We propose to use a ranking technique for feature selection with the help of a completely randomized design method for features and square deviations of property weights with their applications to familiar measures such as syntactic features, vocabulary features and 'most frequent words'. It has been established that besides part of speech probabilities, their entropies play an important role in the discrimination process.*

Keywords: Categorization, Entropy, Features, Rank, Completely randomized design, Square deviations.

1. Introduction

Text classification research has been mostly focused on subject or prepositional content of text. As text-based applications become more varied and the amount of information increases tremendously, different aspects of text can be useful for various purposes including classification. We focus on text genre or the style of text that often characterizes the purpose for which the text has been written. Examples for genre are: research article, novel, poem, news article, editorial, homepage, advertisement, manual, court decision etc.

Text categorization is the task of automatically sorting a set of documents into categories from a predefined set. This task, that falls at the crossroads of information retrieval and machine learning, has witnessed a booming interest in last few years due to its applications in automated indexing of scientific articles according to predefined thesauri terms, filing patents into patent directories, selective dissemination of information to information consumers, automated population of hierarchical catalogues of web resources, spam filtering, identification of document genre, authorship attribution, survey coding, and even automated essay grading. Classificatory tasks in computational

[a] *Dept. of Mathematics, University of Kumaun, S. S. J. Campus Almora, INDIA*
E-mail: h.lata.pande@gmail.com, pande_hemlata1011@yahoo.com
[b] *Dept. of Mathematics,University of Kumaun, S. S. J. Campus Almora, INDIA*
E-mail: profdhami@rediffmail.com

linguistics are performed on the basis of text genre detection and authorship attribution. Genre determination of text usually refers to identification of the kind of the text and is very useful in many natural language processing tasks owing to the fact that the approaches to text genre detection are focused on the use of simple measures that can be easily detected and reliably counted, as discussed in the works of Kessler et al (1997). Parsing accuracy and POS tagging accuracy can be increased by taking genre into account. Word senses are also related to the genre of the text, and in information retrieval genre could help to give better search results. A lot of work has been done in text discrimination with the help of different statistical measures. In a comparative analysis of written and spoken genres of English, Biber (1988) has listed an array of text features. He has also outlined five dimensions which can be used to characterize genre (Biber 1989). Karlgren and Cutting (1994) have taken a set of properties of these distinguishing dimensions and has used them to compute a score to categorize texts in pre defined text genre categories. Lee and Myaeng (2004) have used a deviation formula and a discrimination formula of statistical features for feature extraction from training documents. They have also tested the viability of using genre classification information in the automatic classification of documents into subject areas and have concluded that the knowledge of the genre of a given document helps in correct classification. Kelih et al (2005) have tested impact of word length for genres and authors by taking Russian letters and poems written by three different authors and have found that word length is characteristic of genre rather than of authorship. They (Kelih et al, 2006) have also tested the impact of sentence length in text typology. Stamatatos et al (2000a) have used the frequencies of occurrence of the most frequent words as discriminators of text genre and have tested the role of frequencies of occurrence of the most common punctuation marks for accurate text categorization. Stamatatos et al in another work (2000b) have used multiple regression analysis and discrimination analysis for 22 vectors of style marker of token level, syntax level and analysis level for text categorization in terms of genres and also in terms of authors. They have observed that the size of the training set plays a significant role in the results and have also calculated the contribution of each level style markers to the classification procedure. Wolters and Kirsten (1999) have explored the impact of the inclusion of function word frequencies and POS frequencies in genre classification. Klavans and Kan (1998) have focused on the determination of the role of verbs in document analysis by taking verb profile as one of many factors in the determination of text type. Borja and Chappas (2006) have compared the style of various forms of media texts using measures of vocabulary richness and have also used three parameters of Sichel distribution for this purpose. Antic et al (2006) have studied the suitability of parameters of two parametric Poisson models in discriminating text sorts. Stadlober and Djuzelic(2006) have applied multivariate statistical methods for text classification by text length and by moments of the distribution of word length.

In the paper, we have determined the results for the classification process for the four kinds of articles, by taking three types of features separately. The three types of features have been combined in order to obtain better results for the process. Ranking technique has been selected for inferring about the order of the features, that is, which feature has to be given preference over others? The process has been to prefer the features with the lower rank over the features with higher rank so that better results for the classification process could be obtained. The section 3 of the paper introduces the concept of entropies of POS (Parts of speech) and tests their role in discrimination of texts in various categories.

Classification results for three different types of features - Our analysis of 95 categorized test articles (defined later) by discriminant analysis with 3 different sets of vectors, 11 parts of speech frequencies, 5 vocabulary richness measures and probabilities of 48 most frequent words, revealed that the average of calculated F-values ($F = \frac{1}{\frac{\alpha}{P} + \frac{(1-\alpha)}{R}}$ for α = 0.5, where P and R are precision and recall respectively) for different categories are only 0.59, 0.58, 0.67 respectively, whereas the desired values are closer to 1. In order to improve the results for better categorization of texts, we have selected a combination of features from parts of speech properties, from vocabulary measures and from the most frequent words.

2. Method

For the current study, 188 articles, 96 press news reportages, 44 press editorials, 23 creative writings (novels and short stories) and 25 recipes have been taken from the ICE India Corpus[c] compiled by Shivaji University, Kolhapur, India, and by Freie Universität Berlin, Germany. 49% of all the articles (48 reportages, 22 editorials, 11 creative writings and 12 recipes) were taken as training articles and the remaining (48 reportages, 22 editorials, 12 creative writings and 13 recipes) have been used for the testing purposes. For categorizing the 95 test articles we have used the linear discriminant analysis (with the help of statistiXL[d]) along the lines of Karlgren and Cutting (1994). The results of this process depend on selected features because all features are not equally effective in the discrimination process. Three types of features were used: POS features, vocabulary features and most frequent words in the following manner-

i) Eleven part of speech features: namely frequencies of occurrence of noun, pronoun, adjective, adverb, verb, determiner, conjunction/preposition, model, *to*, cardinal number and interjection have been selected by tagging the texts with the Go tagger[e].

ii) Five vocabulary richness measures propounded by Yule, Honore, Brunet, Sichel, and Simpson as cited in Stamatatos et al, (2000b) have been used, where V is vocabulary of a text N is total word tokens in the text and V_i is number of words i times

$$\left.\begin{aligned} &K = \frac{10^4\left(\sum_{i=1}^{\infty} i^2 V_i - N\right)}{N^2},\ R = \frac{(100 \log N)}{\left(1 - \left(\frac{V_1}{V}\right)\right)},\ W = N^{V^{-\alpha}}, \\ &S = \frac{V_2}{V},\ D = \sum_{i=1}^{V} V_i \frac{i(i-1)}{N(N-1)} \end{aligned}\right\} \quad \text{.....(2.1)}$$

iii) Forty eight most frequent words of English from the 50-words list of Stamatatos et al (2000a) have been selected except *n't* and *'s*.

The calculated values of precision and recall, defined in equation (2.2), for the above three cases have been enumerated in the table1.

---

[c] http://www.ucl.ac.uk/english-usage/ice/index.htm.

[d] http://www.statistixl.com.

[e] http://uluru.lang.osaka-u.ac.jp/~k-goto/download/download.cgi?name=GoTagger07.zip

Table 1. Values of precision and recall for categorization of text articles by POS frequencies, vocabulary measures and 48 most frequent words

Category	11 POS frequencies		Vocabulary measures		48 most frequent words	
	Precision	Recall	Precision	Recall	Precision	Recall
1.Recipes	0.455	0.385	0.529	0.692	0.765	1.0
2.Creative-writing	0.889	0.667	0.5	1.0	0.778	0.583
3.Editorials	0.55	0.5	0.5	0.773	0.444	0.545
4.Reportages	0.618	0.708	0.75	0.3125	0.691	0.604
Average Value:	0.628	0.565	0.569	0.694	0.669	0.683

$$\left.\begin{aligned} precision &= \frac{|\{relevant\ \ documents\ \}\cap\{retrieved\ \ documents\ \}|}{|retrieved\ \ documents\ |} \\ recall &= \frac{|\{relevant\ \ documents\ \}\cap\{retrieved\ \ documents\ \}|}{|relevant\ \ documents\ |} \end{aligned}\right\} \quad .....(2.2)$$

Table 1 depicts that by using the three types of features, the values of precision and recall are not very fine. We have made an attempt in this paper to apply the concept of entropy, and of ranks for features in order to improve text categorization results. Ranking Technique is a way of gathering information on a specific subject, and ranking the importance of the information, in order to meet some decision-making objective. e.g., by gathering the information for the data of the feature values corresponding to various features for the considered training articles (the set of articles, each labeled with their kind and used for extraction of some statistical measures corresponding to some features of texts) to determine - what feature has the highest priority for the categorization of these (training) articles as well as the remaining set of articles called the testing articles.

3. Selection of Features and Role of Entropies of POS

We have chosen a combination of three different types of features by applying filters of completely randomized design and 'square deviations from the mean of feature weights', since the number of features was large.

Assuming recipes, creative-writings, editorials and reportages as first, second, third and fourth categories respectively, we have calculated the differences in means of feature values for each category with respect to the other categories by applying Completely Randomised Design with hypothesis -

H_0: the means of different treatments do not differ significantly and

H_1 : there is a significant difference in means,

For the purpose we have assumed 4 different treatments corresponding to determination of the weights of features for training articles of 4 categories, that is, assuming category 1 for recipes, 2 for creative writings, 3 for press editorials and 4 for press news reportages, treatment 1 has been taken for the determination of weights for the training articles of recipe and similarly for the other categories. If a selected feature *f* has significant difference in means for any one of the four categories with respect to the other categories, we have denoted denote this as $<(i,j)>_f$ (where *i* and *j* represent any two categories having significance difference in their mean values for the property *f* , *i* and *j* both varying from 1 to 4).

For better results in categorization of test texts besides significant difference in means, it is also necessary that various texts of a particular category take feature values which do not differ very much from the mean value of the feature for that category. It shall enable the comparison of means of different categories as comparison of the texts of the different categories and thus the sum of square deviations of feature values of objects from a particular category from their respective means should be lower. But as for different features, feature values are in different ranges so for comparison of their square deviations we have taken weight of a feature in such a way that it is independent of scale.

Defining the weight of a feature as the feature value divided by its maximum value in 93 training texts, we have

$$w_{pq} = \frac{|f_{pq}|}{\max_{1 \le p \le 4}(|f_{pq}|)} \quad q = 1,2,3,....n_p \quad .....(3.1)$$

where f_{pq} *represents feature value of* q^{th} *object of* p^{th} *category* n_p *is no. of training texts in* p^{th} *category*.

Since the square deviation increases with the number of objects, we have computed the values for the resultant variance S_{ij}^2 corresponding to each pair *(i, j)* of categories with significance difference in their means of the feature by the following formula:

$$s_{ij}^2 = \frac{\sum_{k=1}^{n_i} (w_{ik} - \overline{w}_i)^2 + \sum_{k=1}^{n_j} (w_{jk} - \overline{w}_j)^2}{n_i + n_j - 2} ..... \quad .....(3.2),$$

n_i is the number of objects of the i^{th} category.

The average value of S_{ij}^2 has been determined for all discriminating pairs, e.g. if the means of a feature *f* differ significantly for the first and second category and for second and fourth category only, then we have denoted this as $<(1,2),(2,4)>_f$ and have determined the values of s_{12}^2 and s_{24}^2 from equation (3.2) and then have calculated their average value as : Average(s_{ij}^2) = average of s_{12}^2 and s_{24}^2. The process has been thrashed out in detail for the property 'occurrence of noun in text' as-

The frequency of occurrence of noun in a text has been determined for all the considered texts and feature has been taken 'probabilities of occurrence of noun in the text'. We have obtained the weights for this feature, as defined by the equation (3.1) for the 4 categories and the determined values have been shown in Table2.

Table2

Treat-ments		Weights.(w_{pq})
1	Category1	0.685, 0.637, 0.706, 0.685, 0.667, 0.733, 0.640, 0.692, 0.767, 0.851, 0.827, 0.833
2	Category2	0.479, 0.416, 0.385,0.377, 0.363, 0.390, 0.402, 0.316, 0.457, 0.413, 0.295
3	Category3	0.513, 0.534, 0.517, 0.491, 0.513, 0.508, 0.395, 0.554, 0.481, 0.465, 0.564, 0.500, 0.554, 0.539, 0.431, 0.464, 0.463, 0.489, 0.468, 0.494, 0.526, 0.472
4	Category4	0.543, 0.637, 0.607, 0.833, 0.679, 0.733, 0.691, 0.627, 0.710, 0.680, 0.427, 0.687, 0.743, 0.668, 0.619, 0.535, 0.594, 0.574, 0.688, 0.558, 0.679, 0.489, 0.454, 0.975, 0.690, 0.845, 0.693, 0.685, 0.871, 0.877, 0.748, 0.573, 0.556, 0.808, 0.717, 0.520, 0.741, 0.577, 0.769, 0.635, 0.616, 0.718, 0.507, 0.717, 0.964, 0.643, 0.648, 1.000

If $\sum_j y_{ij} = T_i$, be the observation total of i^{th} treatment which is replicated r_i times(i =1, 2, 3, 4), then we have

T_1 = 8.724, T_2 = 4.293, T_3 =10.937, T_4 = 32.546, r_1 =12, r_2 =11, r_3 =22, r_4 =48.
Therefore

Treatment sum of squares $= \sum_{i=1}^{4} \frac{T_i^2}{r_i} - \frac{(\sum_{i=1}^{4} T_i)^2}{\sum_{i=1}^{4} r_i} = 1.197,$

Total sum of squares $= \sum_{i,j} y_{ij}^{\ 2} - \frac{(\sum_{i=1}^{4} T_i)^2}{\sum_{i=1}^{4} r_i} = 2.106,$

Thus the analysis of variance table can be designed in the form of Table 3.

Since the tabulated value of F at 0.05 level of significance is less than the calculated value, we can reject H_0 at 5% level of significance which implies that the means of at least weights of two categories differ significantly at 5% level.

Table3. Analysis of variance table.

Source of variation	Degrees of freedom (d.f.)	Sum of squares	Mean sum of squares	F
Treatments	3	1.197	0.399	39.06
Error (within treatments)	89	0.909	0.0102	
Total	92			

For pairs of treatments (i, j) with significant differences in their means, we should have:

$$\left|\mu_i - \mu_j\right| > t_{0.05} \ \ for \ \ error \ \ \ d.f. \times \sqrt{S_E^2\left(\frac{1}{r_i}+\frac{1}{r_j}\right)},$$

μ_i *be the mean of* i^{th} *treatment* .

Thus probability of 'noun' can be differentiated among 5 pairs out of 6 pairs of categories $<(1,2), (1,3), (2,3), (2,4), (3,4)>_{Noun\ probability.}$

Calculated values of S_{ij}^2 for weights of nouns for above categories are:

$s_{12}^2 = 0.00437$, $s_{13}^2 = 0.00309$, $s_{13}^2 = 0.00211$, $s_{24}^2 = 0.01422$, $s_{34}^2 = 0.0120$,

the average value of these, Average(S_{ij}^2) is 0.00716. Similarly the average value of S_{ij}^2 for all features that have significance difference in means for at least two categories has been obtained.

We know that any feature can be considered as a better discriminator among all other comparing features if it differs significantly in all the 6 pairs and also has lower value of average standard deviation as compared to other features. To meet this requirement, we have ranked features in two ways:

- by number of pairs which the feature can discriminate and
- by value of Average(S_{ij}^2)

We determined the average value of ranks assigned by the above-mentioned two ways and arranged features first by ascending order of their average ranks. If the averages ranks were the same then we arranged the features in ascending order on the basis of their ranks obtained by first way. Where these values were also the same, we formed the order on the basis of their average (S_{ij}^2). This process was applied to different part of speech features, vocabulary features and most frequent words in the following manner:

(1) Part of Speech features: For these features we tabulated 28 parts of speech tags (S. No. 1-28) out of 36 parts of speech tags (excluding /FW, /LS, /POS, /SYM, /NN, /NNPS, /NNP, /NNS) as assigned by the Go tagger (In order to reduce the effect of capitalized words we considered only nouns and not its different sub-categories) while the other seven parts of speech (S. No. 29-35) were calculated by the assigned tags. The probability of each part of speech was calculated by assuming 29 POS in all (S. No. 1-29) as all words must be assigned to one of these 29 parts of speech (assigned as a noun or any POS from S. no. 1-28).

Table 4. Part of speeches and used tags.

S. No.	Part of Speech	S. No.	Part of Speech	S. No.	Part of Speech
1	Coordinating conjunction (/CC)	13	wh-determiner (/WDT)	25	Pre-determiner (/PDT)
2	"Adjective, comparative" (/JJR)	14	wh-adverb (/WRB)	26	Particle (/RP)
3	Possessive pronoun (/PRP$)	15	Non wh-determiner (/DT)	27	Modal (/MD)
4	"Adverb, superlative" (/RBS)	16	Adjective, non superlative & non comparative (/JJ)	28	To (/TO)
5	"Verb, gerund/present participle" (/VBG)	17	Personal pronoun (/PRP)	29	Noun (/NNPS+/NNP+/NNS+/NN)
6	"Verb, non 3rd ps. sing. Present" (/VBZ)	18	"Adverb, comparative" (/RBR)	30	Pronoun (/PRP$+PRP+WP$+WP)
7	Possessive wh-pronoun (/WP$)	19	"Verb, past tense" (/VBD)	31	Adjective (/JJR+/JJS+/JJ)
8	Preposition/subord. Conjunction (/IN)	20	"Verb, non-3rd ps. sing." (/VBP)	32	Adverb (/RBR+/RBS+/WRB+/RB)
9	"Adjective, superlative" (/JJS)	21	wh-pronoun (/WP)	33	Verb (/VB+/VBD+/VBG+/VBN+/VBP+VBZ)
10	Adverb [non wh, non superlative & non comparative] (/RB)	22	Existential there (/EX)	34	Determiner (/DT+/WDT)
11	"Verb, base form" (/VB)	23	Interjection (/UH)	35	Preposition/conjunction (/CC+/IN)
12	"Verb, past participle" (/VBN)	24	Cardinal number (/CD)		

Garcia et al (2006) described a new approach to text categorization where the focus is on the amount of information (entropy) in the text. We have utilized the concept of entropy by considering all the sub-categories of parts of speech mentioned in above table from numbers 1-29 and by considering only 14 categories of parts of speech (numbers 22-35) and have represented them by en_1 and en_2 respectively. Thus

$$en_1 = -\sum_i p_i \log_2 p_i \,, i = 1,2.....29,$$ p_i be the probability of i^{th} POS in text by assuming 29 categories of POS, (S. No. 1-29 in Table 4.)

$$en_2 = -\sum_i p_i \log_2 p_i \,, i = 1,2.....14,$$ p_i be the probability of i^{th} POS in text by assuming 14 categories of POS, (S. No. 22-35 in Table 4.)

Entropies of pronouns were calculated by considering four sub-categories of pronouns, namely possessive pronoun, personal pronoun, wh-pronoun and possessive wh-pronoun, of adjective from its three sub-categories, adverb from 4 sub-categories, verb from 6 sub-categories, determiner from 2 sub-categories and preposition/conjunction from 2 categories. The entropy of that part of speech has been taken equal to zero for which no sub-categories have been considered as is the case of noun.

If a POS with non-zero entropy has n subcategories then its maximum entropy will be $n \log_2 n$ and therefore relative entropy = $\frac{entropy}{n \log_2 n}$.

For each POS with non-zero entropy, relative entropy was calculated and a new feature 'sum of entropies of POS' denoted by 'sum' is defined as:

sum= $\sum_i (entropy)_i$ ……..(3.3) , where $(entropy)_i$ is relative entropy of i^{th} POS.

The suitability of these entropic measures can be visualized in the discrimination of different categories. We can demonstrate this fact with the help of following figure in which the plotted values of en_1 for 93 training texts depict significant differences in the three categories of recipes, editorials and creative-writings. The figure also demonstrates the applicability of this measure in the measurement of the discrimination between creative-writings and reportages.

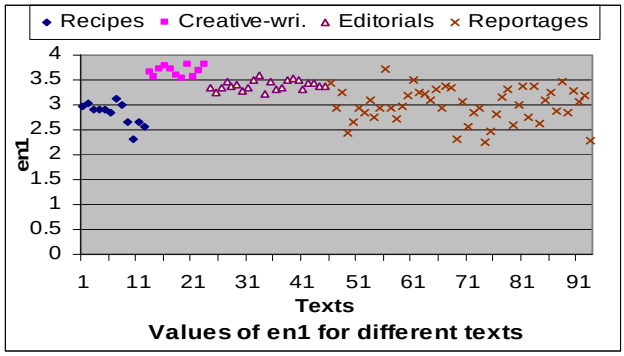


Figure1. Calculated values for training texts.

On the basis of training articles of three types (recipes, creative-writings, editorials), we can consider the numerical values in the following manner:

For creative-writings $en_1 > 3.54$, in case of editorials $3.12 < en_1 < 3.54$, and for recipes $en_1 < 3.12$. We have then classified texts of these three categories on the base of the values of en_1, as shown in the following classification table for the test articles of the three types:

Table5. Classification of testing articles among three categories based on en_1 value.

	Recipes	Creative-writings	Editorials	Total
Recipes	13	0	0	13
Creative-writings	0	10	2	12
Editorials	1	1	20	22

The table exhibits that the classification exercise is 100% for recipes, 91% for Editorials and 83.3% for creative writings.

In a similar manner the separation process between reportages and creative-writings can be classified by taking: $en_1 < 3.5146$ for reportages, which results in correct classification of 11 creative writings out of 12 and 44 reportages out of 48.

Application of this approach in the selection of ranges of en_2 among recipes, creative-writings and editorials under the conditions- for creative-writings $en_2 > 2.861$, for editorials $2.63 < en_2 < 2.861$, and for recipes $en_2 < 2.63$ and also the condition $en_2 < 2.85$ for creative-writings between reportages and creative writings yielded more than 92% average accuracy rate for the three categories of creative-writings, recipes, editorials and also for discrimination of reportages from creative writings in training texts. And respectively had a discrimination of 100% for recipes, 83.3% for creative-writings, 72.7% for editorials correctly among three categories and an average 88.5% between creative writings and reportages for the test texts. The feature 'sum' described by equation (3.3) was used in separation of sets of categories in case of editorials from recipes and creative-writings from recipes.

(2) Vocabulary measures: The five measures under this category were obtained with the help of 5 vocabulary richness statistics defined by equation 2.1.

(3) Most frequent words: We selected 47 words from the list of 50 most frequent words (excluding *n't, 's* on account of their meager occurrence and "*to*", which has been already considered as /TO in POS features) mentioned by Stamatatos et al (2000a).

We observed that out of all these 90 features, 69 features differ in at least 2 categories and the values of Average(S_{ij}^2) for these features lie in the range 0.003-0.057.The process of assignment of ranks was conducted in the following two manners:

a) For categories: Ranks 1, 2, 3, 4, 5, 6 were assigned to features which have significant difference in means for 6, 5, 4, 3, 2, 1 pairs of categories respectively.

b) For Average(S_{ij}^2): Ranks 1, 2, 3, 4, 5, 6 were assigned if the value of Average(S_{ij}^2) for the feature lay in ranges 0-.01, .01-.015, .015-.02,.02-.025,.025-.035, 0.35-0.45 respectively while the rank 7 was assigned for higher values of Average(S_{ij}^2). The motive behind this ranking was to assign lower ranks to features with a lower value of Average(S_{ij}^2). The justification behind assigning the highest value of the rank equal to 7 is that ranks depending on categories are only up to 6 and if we consider much higher values of ranks for Average(S_{ij}^2) then the average of the two ranks will be in the same order as that of II rank.

The discussed features generate following table for assignment of ranks.

Table6. Features arranged by their ranks.

	Feature	Average(S_{ij}^2)	Rank by I manner	Rank by II manner	Average Rank
1	Vocabulary richness function of Simpson, D	0.010997	1	2	1.5
2	Vocabulary richness function of Yule, K	0.011012	1	2	1.5
3	Pronouns	0.011607	1	2	1.5
4	Verbs	0.012158	1	2	1.5
5	en_1	0.003395	2	1	1.5
6	Nouns	0.007161	2	1	1.5
7	Subordinate conjunctions or prepositions	0.009721	2	1	1.5
8	Of*	0.019679	1	3	2
9	en_2	0.003486	3	1	2
10	Prepositions or conjunctions	0.009981	3	1	2

Table 6(Continued)					
11	Adverbs	0.020742	1	4	2.5
12	Cardinal numbers	0.015923	2	3	2.5
13	Determiners	0.017864	2	3	2.5
14	Non wh-determiners	0.0189	2	3	2.5
15	Personal pronoun	0.012377	3	2	2.5
16	Vocabulary richness function of Honore, R	0.013688	3	2	2.5
17	Vocabulary richness function of Sichel, S	0.014355	3	2	2.5
18	Vocabulary richness function of Brunet, W	0.006452	4	1	2.5
19	Adverbs which are not either wh-adverbs or superlative or comparative	0.02273	2	4	3
20	sum	0.018079	3	3	3
21	Were*	0.002379	5	1	3
22	The*	0.026484	2	5	3.5
23	Verb, past tense	0.029389	2	5	3.5
24	Possessive pronoun	0.030108	2	5	3.5
25	Verb, gerund/present participle	0.018569	4	3	3.5
26	Would*	0.012267	5	2	3.5
27	Wh-pronouns	0.048334	1	7	4
28	Have*	0.035895	2	6	4
29	That*	0.037309	2	6	4
30	Was*	0.04017	2	6	4
31	Had*	0.044692	2	6	4
32	Coordinating conjunction	0.02787	3	5	4
33	But*	0.029569	3	5	4
34	Non superlative non comparative adjectives	0.030964	3	5	4
35	It*	0.031005	3	5	4
36	What*	0.032422	3	5	4
37	To	0.020219	4	4	4
38	And*	0.020315	4	4	4
39	As*	0.022386	4	4	4
40	All*	0.024072	4	4	4
41	There*	0.017028	5	3	4
42	Verb, base form	0.046489	2	7	4.5
43	Has*	0.036117	3	6	4.5
44	Models	0.037161	3	6	4.5
45	Not*	0.039373	3	6	4.5

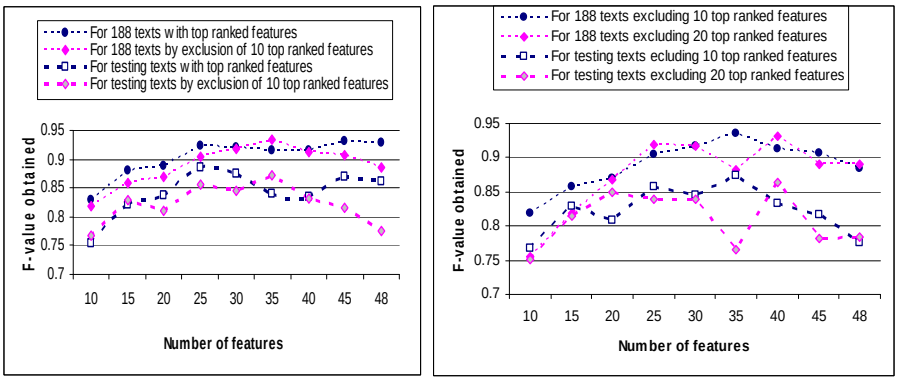
Table 6(Continued)					
46	He*	0.044734	3	6	4.5
47	With*	0.026653	4	5	4.5
48	Verb non 3rd ps. Singular, present	0.027601	4	5	4.5
49	Interjections	0.027821	4	5	4.5
50	Or*	0.031601	4	5	4.5
51	She*	0.032086	4	5	4.5
52	You*	0.033083	4	5	4.5
53	Her*	0.037648	4	6	5
54	Adjectives	0.039904	4	6	5
55	I*	0.041086	4	6	5
56	For*	0.042158	4	6	5
57	Is*	0.044246	4	6	5
58	Wh-determiner	0.025662	5	5	5
59	In*	0.026881	5	5	5
60	Wh-adverbs	0.034098	5	5	5
61	Verb non 3rd ps. Singular	0.04863	4	7	5.5
62	We*	0.050809	4	7	5.5
63	Be*	0.051602	4	7	5.5
64	They*	0.052293	4	7	5.5
65	Been*	0.036836	5	6	5.5
66	By*	0.040053	5	6	5.5
67	On*	0.048059	5	7	6
68	Said*	0.050816	5	7	6
69	Adjectives, comparative	0.057112	5	7	6

In this case for each part of speech, the feature has been selected as its probability of occurrence and for frequent words (marked by * in table) value of 'frequency of word/total word tokens' has been used. This table envisages that all the three entropies, defined by us, have obtained positions in top 20 ranked features (two entropies, en_1 and en_2 takes position in top 10 ranked features) out of all the 69 features this implies that entropies of texts are good features (as getting lower ranks) for text categorization and therefore can be utilized as better measures for genre discrimination.

4. Observations and Results

In this section we will argue for the advantage of using the investigated method for the feature selection process. For this purpose we have applied discriminant analysis to all 188 (training as well as testing). We took the different features and observed that the average values of F when no ranks are excluded with first 10, 15, 20, 25, 30, 35, 40, 45, 48 features of Table5 are .828, .879, .888, .923, .921, .914, .915, .932, .928; when top 10 ranks are excluded .819, .858, .87, .905, .918, .935, .914, .907, .885 and when 20

ranks are excluded are .755, .818, .868, .918, .918, .882, .931, .889, .891 respectively. In a similar manner we have obtained the average values for the testing texts under the same conditions. The results are graphically represented in the following figures:


For 188 texts with top ranked features
For 188 texts by exclusion of 10 top ranked features
For testing texts with top ranked features
For testing texts by exclusion of 10 top ranked features
F-value obtained
0.95
0.9
0.85
0.8
0.75
0.7
10 15 20 25 30 35 40 45 48
Number of features
For 188 texts excluding 10 top ranked features
For 188 texts excluding 20 top ranked features
For testing texts ecluding 10 top ranked features
For testing texts excluding 20 top ranked features
F-value obtained
0.95
0.9
0.85
0.8
0.75
0.7
10 15 20 25 30 35 40 45 48
Number of features

Figures 2 and 3. F values obtained for all articles and also for testing articles by taking different number of features.

From the above figures it is clear that the highest average value of F for the testing articles is obtained as 0.89 by using 25 top ranked features while it is 0.87 when 10 top ranked features are excluded and 35 features are used. In the case of the exclusion of 20 top ranked features, the maximum F value obtained was to be 0.86 in 40 features for the testing texts. These accentuate the inclusion of top ranked features for categorization of testing articles for getting better results.

We are now in a position to establish the superiority of the discussed methods over the prevalent methods. We have constructed the classification table for the testing articles corresponding to highest value of F, which is shown in the following table:

Table7. Classification table for testing texts with 25 top ranked features.

Predicted Category / Actual Category	Recipe	Creative-writing	Editorial	Reportage	Total	Precision / Recall
Recipe	13	0	0	0	13	0.87/1
Creative-writing	0	12	0	0	12	0.92/1
Editorial	0	0	20	2	22	0.71/0.91
Reportage	2	1	8	37	48	0.95/0.77
Total	15	13	28	39	95	

For this classification table the average values of precision and recall are 0.92 and 0.86 respectively, which are better than the values obtained with three types of features shown in table1. Repetition of the process with a slight change of assignment of

second rank as 1,2,3,4,5,6 for Average(S_{ij}^{2}) in ranges 0-.01,.01-.015,.015-.02,.02-.025,.025-.03 or higher reveals that the average value of F for the testing articles is maximum (which is 0.88) for top ranked features while the maximum of average F value with the exclusion of 10 top ranked and 20 top ranked features are 0.87 and 0.86 respectively.

The above discussion signifies the fact that the ranking technique assigns lower ranks to features, which in turn will give better results for the categorization of the testing articles. This technique enables better feature selection from training texts so that results for the testing texts can also be improved.

5. Conclusion

On the basis of the discussions and observations carried out in the present paper, it can be concluded that the concept of entropies of POS are good features for the discrimination of texts. The suitability of these entropic measures has been visualized in the discrimination of different categories.

We can also conclude that while applying the discrimination process, instead of taking some particular kinds of features separately as POS features, vocabulary measures and most frequent words-list we can take a set of features comprising all features by applying filters of completely randomized design and square deviations from the mean of feature weights which have a possibility to give better results as the results of the process depend on the features taken. This possibility has been expressed by us in the form of a ranking technique so that the features with lower ranks should be taken preferentially than the features with higher ranks. This technique shall enable better feature extractions from training texts so that results for the testing texts can also be improved.

Acknowledgement

The authors are grateful to the Council of Scientific & Industrial Research (CSIR), New Delhi, for providing financial assistance to carry out the research work on this interesting field of applied Mathematics in the form of a senior research fellowship to the first author (CSIR grant number : 09/428(0062)/2006-EMR-1).

References

Antic G.; Stadlober E.; Grzybek P.; Kelih E. (2006). Word length and frequency distributions in different text genres. In M. Spiliopoulou et al. (Eds.): From Data and Information Analysis to Knowledge Engineering, pp. 310-317, Springer, Heidelberg.

Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press.

Biber, D. (1989). A typology of English texts. Language, **27** , pp. 3-43.

Borja, M., C.; Chappas, C. (2006). A Stylometric Analysis of Newspapers, Periodicals and News Scripts. Journal of Quantitative Linguistics, **13**(2-3), pp. 285-312.

Garcia, M.; Hidalgo, H.; Chavez, E. (2006). Contextual Entropy and Text Categorization. Fourth Latin American Web Congress (LA-WEB'06), pp. 147-153.

Karlgren, J.; Cutting, D. (1994). Recognizing text genres with simple metric using discriminant analysis. Proceedings of the 15th Conference on Computational Linguistics, Vol. 2, pp. 1071–1075.

Kelih, E.; Antic, G.; Grzybek, P.; Stadlober, E. (2005). Classification of Author and/or Genre? The Impact of Word Length. In: C. Weihs, and W. Gaul (Eds.), Classification – The Ubiquitous Challenge. Springer, Berlin, pp. 498–505.

Kelih, E.; Grzybek, P.; Antic, G.; Stadlober, E. (2006). Quantitative Text Typology: The Impact of Sentence Length. In: M. Spiliopoulou et al. (Eds.): From Data and Information Analysis to Knowledge Engineering.Springer, Berlin, pp. 382–389.

Kessler, B.; Nunberg, G.; Schuetze, H. (1997). Automatic detection of text genre. In Proceedings of the 35th Annual Meeting on Association for Computational Linguistics, pp. 32–38.

Klavans, J.; Kan, M., Y. (1998). Role of verbs in document analysis. *In proc. COLING/ACL* Montreal.

Lee, Y.; Myaeng, S. (2004). Automatic Identification of Text Genres and Their Roles in Subject-Based Categorization, Proc. 37 HICSS, 2004.

Stadlober, E.; Djuzelic, M. (2006). Multivariate Statistic Methods of Quantitative Text Analysis. In: Grzybek, Peter (ed.), Contributions to the Science of Text and Language. Word Length Studies and Related Issues.Dordrecht, NL: Springer; pp. 259-276.

Stamatatos, E.; Fakotakis, N.; Kokkinakis, G. (2000a). Text genre detection using common word frequencies. In 18th International Conference on computational Linguistics, 2000.

Stamatatos, E.; Fakotakis, N.; Kokkinakis, G. (2000b). Automatic text categorization in terms of genre and author. Computational Linguistics, **26**(4), pp. 471–495.

Wolters, M.; Kirsten , M. (1999). Exploring the use of linguistic features in domain and genre classification. In *Proceedings of EACL'99*, Morristown, NJ, USA. pp. 142–149.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 1, pp. 49 - 58, 2011

IDENTIFYING AND SEQUENCING OF THE ELEMENTARY CONCEPTS OF MEASUREMENT OF MASS

SK. SAMSUL ALAM
Research Scholar, Centre for Mathematics Education, Visva-Bharati, West Bengal, India-731235.
E-mail: samsulalam_s@yahoo.in

In this paper attempts have been made (a) to identify all the elementary concepts of the major concept 'measurement of mass' and (b) to find the sequential order of these elementary concepts. Total 715 elementary concepts have been identified and sequenced.

Keywords: measurement, mass, elementary concepts.

1. INTRODUCTION:

Mathematics nourishes our intellectual faculty. It improves our habits of argumentation, power of reasoning, judgment of any difficult situation and arriving at a correct decision. These are the characteristics of a civilized person. Thus mathematics plays a silent role in promoting our civilization. But it is observed that in our formal education, many learners can not achieve the specific content of school mathematics properly. They became weak in the subject. In this regard, Becker (2001) said, "Research has documented the failure of mathematics education for too many elementary school students." In India, the present status of school mathematics is not cheering up according to different reports and comments of mathematicians (Buch, 1991, Batra, 1995). But according to the great psychologist & philosopher Piaget (1973), "Every normal child is capable of learning mathematics." Therefore some causes are involved on account of the said unwanted situation. Improper content organization in the syllabi is an important factor that will be discussed in this paper. In this connection, Pushpanadham (1998) stated, "Mathematics is the subject which requires proper understanding of the concepts and their interrelations." Chilana (1984) opined, "Mathematical concepts should be graded in such a way that there is a proper sequence in their growth and development." Again the comment of 'Ben Clarke Pacific Institutes for Research' (2006) on Early Mathematics Assessment is "Emerging evidence suggests difficulty in mathematics is relatively stable over time. The importance of a logical scope and sequence is paramount." So it may be concluded that the proper sequence of 'basic concepts' according to their logical and psychological hierarchy should be maintained in the teaching-learning system of mathematics.

In our country, although the syllabus framing and content organization are made on the basis of experts' opinion but no systematic analytical method is applied in this process. As a result, many gaps are developed in the entire teaching-learning system which is very harmful to the students of mathematics because mathematics is purely a logical science. Therefore, to remove the said conceptual gaps, concept identification and its proper sequencing have an important role. The researcher has identified and sequenced the basic concepts of mathematics through 'text book scanning process' and

'task analysis technique'. In this regard, identification of the basic concepts and their sequencing is a worthwhile enterprise which has not yet been undertaken by others.

It has been made some contribution towards identification and sequencing of some basic concepts and framing of a model syllabus also (Sk., 2003a; 2003b; 2005a; 2005b; 2006a; 2006b; 2006c; 2006d; 2007a; 2007b; 2007c; 2007d; 2008a; 2008b; 2008c; 2008d; 2008e; 2008f; 2008g; 2009a; 2009b; 2009c; 2009d; 2009e; Sk. & Samanta, 2009a) .

In this paper, an attempt has been made for identification and sequencing the elementary concepts of measurement of mass.

2. DEFINITIONS:

2.1 Major concept and Sub-concepts:

A Major concept is an idea which is complete in itself and is comprehended through a sequential process of step-by-step partial comprehension of its related concepts. These related concepts are called sub-concepts which are not complete in itself but are parts of the major concept.

Examples of major concepts are 'addition of whole numbers including the familiarity of the numbers', 'subtraction of whole numbers', 'measurement of mass', 'fraction', 'decimal', 'rational number' etc.

For the major concept such as 'Addition of whole numbers including the familiarity of the numbers', its sub-concepts are 'addition of one-digit numbers including zero where the sum is one-digit number', 'addition of one-digit numbers where the sum is two-digit number' etc.

2.2 First level sub-concepts:

The sub-concepts which are obtained after immediate derivation of a major concept are called first level sub-concepts.

2.3 Second level sub-concepts and others:

When the derivation of first level sub-concepts are continued, the other sub-concepts so obtained stage by stage are called second level, third level etc.

2.4 Elementary concept:

A sub-concept which can not be split further will be called an 'elementary concept'.

3. OBJECTIVE OF THE STUDY:

The objective of the study is to identify different elementary concepts and their sequential order of measurement of mass.

4. PROCEDURE:

Procedure adopted has been discussed below:

4.1 Collection of text books:

The prescribed text books of mathematics from Class-1 to 10 of West Bengal Board of Primary Education (WBBPE), West Bengal Board of Secondary Education (WBBSE), National Council of Educational Research and Training (NCERT) and other available books from the market were collected in the first stage (Malhotra & Gupta, 2000; 2002a; 2002b; 2002c; NCERT, 1996; 1997; 1998; 2000; 2002a; 2002b; 2003a; 2003b; 2004a; 2004b; WBBPE, 2001a; 2001b; 2002a; 2002b; WBBSE, 1998; 2001; 2004; 2005)

4.2 Analysis of text books:

These books were analyzed to identify 'measurement of mass' as a major concept and their elementary concepts of arithmetic mainly.

4.3 Task analysis of question papers:

In the identification and sequencing of basic concepts, task analysis technique has also been applied on arithmetical and algebraic problems of mathematics question papers of 10^{th} standard of Madhyamik Pariksha (1999, 2000) of WBBSE, Indian Certificate of Secondary Education (ICSE) Examination (2005, 2006) of the Council for the Indian School Certificate Examinations, School Certificate Examination (1999, 2000) of Visva-Bharati (a Central University).

4.4 Sequencing of concepts:

The identified major concept and their elementary concepts were sequenced keeping in view the logical order of the subject and the psychological order of learners. The gaps in the concepts detected by the researcher were filled in by him at the initial stage.

4.5 Experts' opinions:

The major concept and their elementary concepts with examples were given to experts for their comments. The experts were requested to add or omit or alter the sequence of concepts as they felt necessary.
Finally, the sequential form of elementary concepts of the major concept 'measurement of mass' incorporating the experts' opinion was developed.

5. SALIENT POINTS OF STUDY:

Total 715 elementary concepts have been identified and sequenced after **analyzing** the major concept 'measurement of mass'. The hierarchy of different levels is made according to the logical order of the concept. At the first level sub-concepts comprise introduction of the concept, units of measurement, conversion of different units and arithmetical operations. Considering this 7 first level sub-concepts are obtained which are expressed in a tabular form in Table-1.

Table-1. First level sub-concepts of the major concept 'Measurement of mass'

Serial nos.	First level sub-concepts
1	Concept of mass and its units
2	Conversion of units of mass
3	Addition involving mass

4	Subtraction involving mass
5	Multiplication involving mass
6	Division involving mass
7	Simplification involving mass

The splitting of sub-concepts is continued through different levels until elementary concepts are reached. The result obtained in such a process is given below in a tabular form.

Table-2. Detailed list of different levels of sub-concepts of the major concept: measurement of mass

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
First level sub-concept sl. nos.	Number of second level sub-concept	Number of third level sub-concept	Number of fourth level sub-concept	Number of fifth level sub-concept	Number of sixth level sub-concept	Number of seventh level sub-concept	Number of eighth level sub-concept	Total number of elementary concepts
1	10	1+1+1+ 1+1+ 1+1+1+1+2	—	—	—	—	—	11
2	2	2+2	5+5	8+8	—	—	—	16
3	2	6+4	12+8	46+32	52+32	—	—	84
4	2	6+4	12+8	46+32	52+32	—	—	84
5	2	6+4	18+12	70+48	76+48	—	—	124
6	2	6+4	18+12	72+48	—	—	—	120
7	2	2+2	4+8	16+40	79+90	182+90	186+90	276
							Grand Total	**715**

The Table-2 depicts the major concept: measurement of mass which has seven (1,2,……,7) first level sub-concepts. In this investigation each sub-concept is numbered with a position value of different levels i.e. when a sub-concept is denoted by i.j.k……., then i indicates the first level sub-concept no., j indicates the second level sub-concept no., k indicates the third level sub-concept no. etc.

Analysis of detailing out of the first level sub-concept no-2 only has been done for the economy of space. Total 16 elementary concepts of sub-concept no-2 have been presented below. As and when an elementary concept is reached it is illustrated by an example with a 'Q' mark.

The second level sub-concepts of sub-concept no. 2: Conversion of units of mass.
2.1: Conversion of units of mass without using decimal system.
2.2: Conversion of units of mass with using decimal system.

The third level sub-concepts of sub-concept no. 2.1: Conversion of units of mass without using decimal system.

2.1.1: Expression of mass involving different units from word form to numeral form and vice-versa.

2.1.2: Changing one unit from another.

The fourth level sub-concepts of sub-concept no. 2.1.1: Expression of mass involving different units from word form to numeral form and vice-versa.

2.1.1.1: Expression of mass involving different units from word form to numeral form.

Q. Express into numeral form.

Four kilograms seven decagrams three grams.

2.1.1.2: Expression of mass involving different units from numeral form to word form.

Q. Express into word form.

Kilogram	hectogram	decagram	gram	decigram	centigram	milligram
2	3	4	5	1	7	0

The fourth level sub-concepts of sub-concept no. 2.1.2: Changing one unit from another.

2.1.2.1 Changing units involving kilogram and gram.

2.1.2.2 Changing units involving gram and milligram.

2.1.2.3 Changing units involving any unit.

The fifth level sub-concepts of sub-concept no. 2.1.2.1: Changing units involving kilogram and gram.

2.1.2.1.1 Changing higher units into lower units.

Q. Convert into gram.

5 kilogram 200 gram

2.1.2.1.2 Changing lower units into higher units (expressed in mixed unit form).

Q. Convert into mixed unit form.

7300 gram.

The fifth level sub-concepts of sub-concept no. 2.1.2.2: Changing units involving gram and milligram

2.1.2.2.1 Changing higher units into lower units.

Q. Convert into milligram.

2 gram 500 milligram

2.1.2.2.2 Changing lower units into higher units (expressed in mixed unit form).

Q. Convert into mixed unit form.

2500 milligram.

The fifth level sub-concepts of sub-concept no. 2.1.2.3: Changing units involving any units.

2.1.2.3.1 Changing higher units into lower units.

Q. Convert into centigram.

4 gram.

2.1.2.3.2 Changing lower units into higher units (expressed in mixed unit form).

Q. Convert into mixed unit form.

222 centigram.

The third level sub-concepts of sub-concept no. 2.2: Conversion of units of mass with using decimal system.

2.2.1: Expression of mass involving different units from word form to numeral form and vice-versa.

2.2.2: Changing one unit from another.

The fourth level sub-concepts of sub-concept no. 2.2.1: Expression of mass involving different units from word form to numeral form and vice-versa.

2.2.1.1: Expression of mass involving different units from word form to numeral form.

Q. Express into numeral form.

Seventy thousand seven hundred point one two grams.

2.2.1.2: Expression of mass involving different units from numeral form to word form.

Q. Express into word form.

Kilogram	hectogram	decagram	gram	decigram	centigram	milligram
	7	0	4	.	5	

The fourth level sub-concepts of sub-concept no. 2.2.2: Changing one unit from another.

2.2.2.1 Changing units involving kilogram and gram.

2.2.2.2 Changing units involving gram and milligram.

2.2.2.3 Changing units involving any unit.

The fifth level sub-concepts of sub-concept no. 2.2.2.1: Changing units involving kilogram and gram.

2.2.2.1.1 Changing higher units into lower units or mixed units.

Q. Express into gram changing the place of decimal point.

210.25 kilograms

2.2.2.1.2 Changing lower units into higher units.

Q. Express into kilogram changing the place of decimal point.

75.5 gram

The fifth level sub-concepts of sub-concept no. 2.2.2.2: Changing units involving gram and milligram.

2.2.2.2.1 Changing higher units into lower units or mixed units.

Q. Express into milligram changing the place of decimal point.

210.25 grams

2.2.2.2.2 Changing lower units into higher units.

Q. Express into gram changing the place of decimal point.

2 milligram

The fifth level sub-concepts of sub-concept no. 2.2.2.3: Changing units involving any units.

2.2.2.3.1 Changing higher units into lower units or mixed units.

Q. Express into centigram changing the place of decimal point.

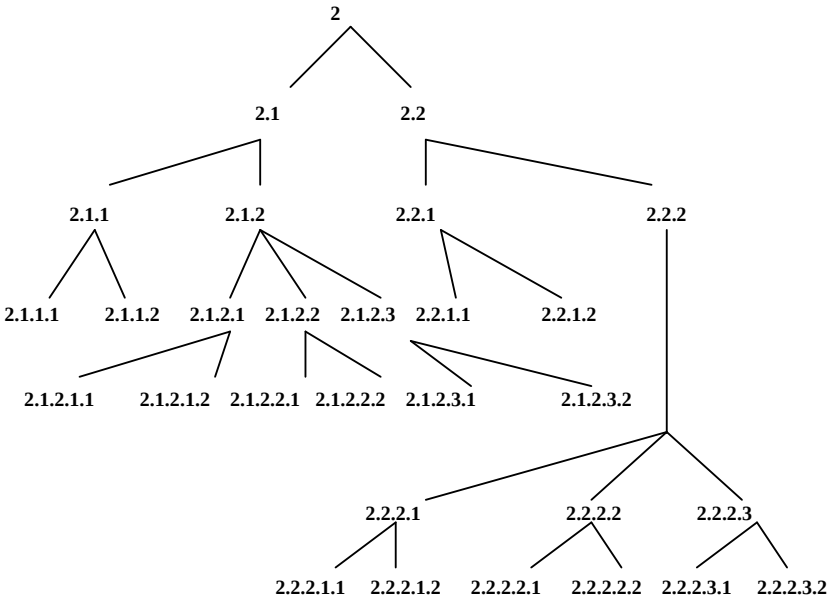
0.004 gram

2.2.2.3.2 Changing lower units into higher units.

Q. Express into gram changing the place of decimal point.

0.5 centigram.

A tree-like structure has been developed from the splitting of the first level sub-concept no.-2 which is depicted below.



6. IMPLIMENTATION:

Ideal gapless and properly sequenced mathematics syllabi for class I to IV standard have been framed (Sk., 2006b; 2008b; 2009e) by implementing the research findings.

7. CONCLUSION:

1. This methodology will help to mark different types of conceptual gaps of mathematics in the syllabus, text books and entire teaching learning process because applying this methodology it has been possible to identify that several conceptual gaps in the existing syllabi and recommended text books of some standard educational Boards. This will help in framing ideal gapless syllabi and developing good text books in future.
2. It will also help to diagnose the particular areas of weakness of students and also in planning for necessary remedial measures.
3. Backward learners can be detected easily and remedial method for them can be applied.

4. Special interest for mathematics can be enhanced which will be helpful for entire science education.

8. FURTHER STUDY:

The diagnostic study may be taken on large number of samples of classes I, II, III & IV on 'measurement of mass'.

ACKNOWLEDGEMENT:

I would like to thank Adhyapak Dr. A. Sensarma, Shiksha-Satra, Visva-Bharati, West Bengal, India and Prof. S. K. Samanta, Department of Mathematics, Visva-Bharati and Professor-in-charge, Centre for Mathematics Education, Visva-Bharati, West Bengal, India. I am also thankful to learned referee and editorial coordinator of the journal for rewriting the paper in the present form.

REFERENCE:

Batra, B. (1995). Learning to use Mathematics at upper K.G. Level, The Primary Teacher, **20(4),** P-10 ,New Delhi, India: NCERT.

Ben Clarke Pacific Institutes for Research in www.centeroninstruction.org, 2006

Buch, M.B. (Ed.) (1991). Fourth Survey of Research in Education, 1983-88, **Vol.-1**, P-24, New Delhi, India: NCERT.

Chilana, R (1984). Teaching Elementary School Mathematics, The Primary Teacher, **9(4),** P-3, New Delhi, India: NCERT.

Malhotra,, O.P. & Gupta, S.K. (2000). Mathematics Today (Revised Edition), Book-3 (for class-3), S. Chand & Company Ltd., Ram Nagar, New Delhi-110055.

_____(2002a). Mathematics Today (Revised Edition), Book-1 (for class-1), S. Chand & Company Ltd., Ram Nagar, New Delhi-110055.

_____(2002b). Mathematics Today (Revised Edition), Book-2 (for class-2), S. Chand & Company Ltd., Ram Nagar, New Delhi-110055.

_____(2002c). Mathematics Today (Revised Edition), Book-4 (for class-4), S. Chand & Company Ltd., Ram Nagar, New Delhi-110055.

National Council of Educational Research and Training (NCERT) (1996). Let's Learn Mathematics (Book One), February, New Delhi, India: NCERT.

_____(1997). Let's Learn Mathematics (Book Three), January, New Delhi, India: NCERT.

_____(1998). Let's Learn Mathematics (Book Four), January, New Delhi, India: NCERT.

_____(2000). Mathematics (with supplement) : A text book for Class-X, January, New Delhi, India: NCERT.

_____(2002a). Mathematics: A text book for Class 9, February, New Delhi, India: NCERT.

_____(2002b). Mathematics: A text book for Class 6, February, New Delhi, India: NCERT.

_____(2003a). Let's Learn Mathematics, Book Two, January, New Delhi, India: NCERT.

_____(2003b). Let's Learn Mathematics, Book Five, December, New Delhi, India: NCERT.

_____(2004a). Mathematics (Part 1): A text book for Class 7, January, New Delhi, India: NCERT.

_____(2004b). Mathematics: A text book for Class 8, January, New Delhi, India: NCERT.

Piaget J. (1973). To Understand is to Invent, Viking, New York.

Pushpanadham. K. (1998). An investigation into the Pupils' Understanding of Basic Arithmetic Concepts, Quest in Education, **22(2),** P-38.

Sk., Samsul Alam (2003a). Sequencing the concepts of subtraction of whole numbers, Journal of Centre for Pedagogical Studies in Mathematics, Thirteenth Issue, PP5 - 13.

_____(2003b). Sequencing the concepts of addition of whole numbers, Indian Journal of Mathematics Teaching, **29(1&2),** PP13 – 23.

_____(2005a). Sequencing the concepts of multiplication of whole numbers, Journal of Centre for Pedagogical Studies in Mathematics, Joint Issue, **14 & 15,** PP16 – 22.
_____(2005b). Sequencing the concepts of division of whole numbers, Indian Journal of Mathematics Teaching, **31(1&2),** PP26 – 37.
_____(2006a). Sequencing of the integral exponent of rational number, Journal of Centre for Pedagogical Studies in Mathematics, Sixteenth Issue, PP15 – 21.
_____(2006b). Identification and Sequencing of the Mathematical concepts of class-I standard, Quest in Education, **30(3),** PP20 – 24.
_____(2006c). Sequencing the concepts of the operations on algebraic expression, Indian Journal of Mathematics Teaching, **32(1&2),** PP. 60-67.
_____(2006d). Karanir dharanagulir dharabahikikaran, Ganit Parikrama, Bingswati Barsha, **1&2,** PP11-19.
_____(2007a). Sequencing the concepts of International numerals sand Roman numerals, Journal of Centre for Pedagogical Studies in Mathematics, Seventeenth Issue, PP. 17-21.
_____(2007b). Sequencing the concepts of factorization of algebraic expression, Vigyan Shikshak, **51(1&2).**
_____(2007c). Sequencing the concepts of rational expression, Indian Journal of Mathematics Teaching, **33(1&2),** PP26-31.
_____(2007d). Asannamaner dharanagulir dharabahikikaran, Ganit Parikrama, Ekbingswati Barsha, **1&2,** PP4-8.
_____(2008a). Sequencing the concepts of decimal, Parshad Barta, Chautrishtam Barsha, January '08, PP 22-32, Paschim Banga Madhya Shiksha Parshad.
_____(2008b). Identification and sequencing of the mathematical concepts of class-II standard, Quest in Education, **32(3),** PP23-30.
_____(2008c). Sequencing the concepts of Quadratic Equation, Journal of Centre for Pedagogical Studies in Mathematics, Eighteenth Issue, PP21-26.
_____(2008d). Sequencing the concepts of ratio (algebraic treatment), Vigyan Shikshak, **53(1&2),** 19-22.
_____(2008e). Swabhabik sankhyar saraler dharanagulir dharabahikikaran (Sequencing the concepts of simplification of whole numbers), Annual Bulletin, published in the occasion of : Nineteenth Annual Conference of CPSM.
_____(2008f). Identification and Sequencing of the Elementary Concepts of Cube and Cube Roots, Indian Journal of Mathematics Teaching, **34(1&2),** P39-47.
_____(2008g). Samay o duratwar dharanagulir dharabahikikaran, Ganit Parikrama, Dwabingswati Barsha, **1&2,** PP7-13.
_____(2009a). Identifying and Sequencing of the Elementary Minor Concepts of the Highest Common Factor (H.C.F.) and the Lowest Common Multiple (L.C.M.) of Numbers, Journal of the Korea Society of Mathematical Education Series D: Research in Mathematics Education, **13(1),** 49-61.
_____(2009b). Identification and Sequencing of the Elementary Minor Concepts of Rational Number, Siksha Chintan, **3**, 55-66(March).
_____(2009c). Sequencing the concepts of linear equation and its graphs, Parshad Barta, Panytrishtam Barsha, March-April, PP 25-32, Paschim Banga Madhya Shiksha Parshad.
_____(2009d). Identifying and Sequencing of the Elementary Concepts of Measurement of Length, Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education, **13(2),** 123-139.
_____(2009e). Identification and sequencing of the mathematical concepts of class-III standard, Quest in Education, **33(3),** PP15-23.
Sk., Samsul Alam & Samanta, Prof. S.K. (2009a). Identifying and Sequencing of the Elementary Concepts of Measurement of Time, accepted in the 43rd Korean National Meeting of Mathematics Education. 16-17 oct'2009 for oral presentation.
West Bengal Board of Primary Education (WBBPE) (1999). Naba Ganit Mukul (Pratham Bhag) December. Kolkata, India: WBBPE.

_____(2001a). Naba Ganit Mukul (Dwitiya Bhag) January. Kolkata, India: WBBPE.
_____(2001b). Naba Ganit Mukul (Chaturtha Bhag)January. Kolkata, India: WBBPE.
_____(2002a). Naba Ganit Mukul (Tritiya Bhag). Kolkata, India: WBBPE.
_____(2002b). Naba Ganit Mukul (Pancham Bhag) January. Kolkata, India: WBBPE.
West Bengal Board of Secondary Education (WBBSE) (1998). Integrated Ganit (Class 9 & 10) **Vol. 1.** Kolkata, India: WBBSE.
_____(2001). Mathematics for Class 8, Kolkata, India: WBBSE.
_____(2004). Mathematics for Class 6, May. Kolkata, India: WBBSE.
_____(2005). Mathematics for Class 7, January. Kolkata, India: WBBSE.

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 1, pp. 59- 72, 2011

THE CENTROID OF SOME GENERALIZED PEDAL CONFIGURATIONS

SZILÁRD ANDRÁS*
Department of Applied Mathematics, Babeş-Bolyai University, Cluj Napoca, RO-400084, Romania andraszk@yahoo.com
http://www.math.ubbcluj.ro/~andrasz

The main goal of this paper is to give possible generalizations, analogues of the following property:

If M_1, M_2 and M_3 are the orthogonal projections of a point M to the sides A_1A_2, A_2A_3 and A_3A_1 of an equilateral triangle $A_1A_2A_3$, then the centroid of the triangle $M_1M_2M_3$ is the midpoint of the segment OM, where O is the center of the triangle $A_1A_2A_3$.

In the first step we extend this property to regular n-gons, regular tetrahedrons and regular n simplices. In the second part we give a general affine version for triangles and simplices.

It is also our objective to analyze the possibility of using such properties in teaching problem solving strategies for students and mathematics teachers. Theorem 2, 2.3, 2.5, 2.10 and Conjecture 1 and 2 were discovered/rediscovered during a training course for mathematics teachers.

Keywords: pedal polygons; pedal simplices; centroids.

1. Introduction

In this note we analyze some possible generalizations of the following property:

Theorem 1.1. *If we denote by M_1, M_2 and M_3 the orthogonal projections of a point M to the sides A_1A_2, A_2A_3 and A_3A_1 of an equilateral triangle $A_1A_2A_3$, then the centroid G of the triangle $M_1M_2M_3$ is the midpoint of the line segment OM, where O is the center of triangle $A_1A_2A_3$.*

This problem was studied on a teacher training course at the Babeş-Bolyai University from the perspective of the inquiry based learning. The starting point was not only the problem, but also a solution of this problem using complex numbers. The aim of the instructors was to support the participants (secondary school mathematics teachers) in experimenting, developing and proving generalizations of this property. This approach was used in order to prove that inquiry based learning

*RO-407280, Floreşti, Str. Eroilor, 246, Bl. 4D, Ap. 2, Romania

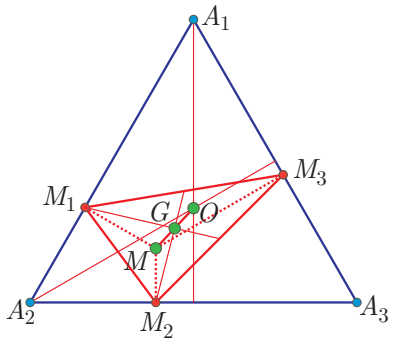


Fig. 1. The centroid of a podar triangle relative to an equilateral triangle

can also be used in problem solving activities and in the framework of the existing curricula. The main tools for experimenting were dynamic geometry softwares (Geogebra, Geonext and Cabri) and the teachers were working in groups. In the first round each group had to present some ideas about the possible generalizations, in the second round each group had to formulate conjectures and had to experiment his own conjectures and in the final round they had to prove the conjectures which seemed to be valid.

2. Conjectures and proofs

As a first step we can replace the equilateral triangle with something more general: an arbitrary triangle, a regular polygon, a regular tetrahedron or a regular simplex. Or we can use some more general projections in constructing the pedal triangle. In order to formulate some more general properties we recall a proof for theorem 1.1.

Proof. We assume M is in the plane of the triangle $A_1A_2A_3$, and the circumscribed circle is of radius 1. Let's consider O as the origin of the coordinate system and OA_1 as the OX axis. The complex numbers corresponding to the vertices A_1, A_2 and A_3 are

$$a_1 = \varepsilon, \quad a_2 = \varepsilon^2 \text{ and } a_3 = \varepsilon^3,$$

where $\varepsilon^3 = 1$ and $\varepsilon \neq 1$. Due to our assumptions, m_1 is easy to find, because it's real part is $-\frac{1}{2}$ and it's imaginary part is the same as m's. So

$$m_1 = -\frac{1}{2} + \frac{m - \overline{m}}{2}.$$

To calculate m_2 we'll use a counterclockwise rotation of angle $\alpha = \frac{4\pi}{3}$. This rotation transforms the point M_2 to the orthogonal projection of the point $Q(\varepsilon^2 \cdot m)$ to the side A_1A_2. Hence we have

$$m_2 \cdot \varepsilon^2 = -\frac{1}{2} + \frac{\varepsilon^2 \cdot m - \left(\overline{\varepsilon^2 \cdot m}\right)}{2},$$

and so

$$m_2 = -\frac{1}{2} \cdot \varepsilon + \frac{m - \varepsilon^2 \cdot \overline{m}}{2}.$$

By the same argument we deduce

$$m_3 = -\frac{1}{2} \cdot \varepsilon^2 + \frac{m - \varepsilon \cdot \overline{m}}{2}.$$

From these relations we get

$$\frac{m_1 + m_2 + m_3}{3} = \frac{m}{2},$$

which expresses the desired property. □

The above presented proof suggests that something similar holds also for a regular n-gon. A little experience with a dynamic geometry software helps us to formulate the following property:

Theorem 2.1. *If we denote by M_i $(1 \leq i \leq n)$ the orthogonal projections of a point M to the sides A_iA_{i+1} $(1 \leq i \leq n$ and $A_{n+1} = A_1)$ of a regular n-gon $A_1A_2A_3 \ldots A_n$ and by O the center of the polygon, then the centroid of the n-gon $M_1M_1M_2 \ldots M_n$ is the midpoint of the line segment OM.*

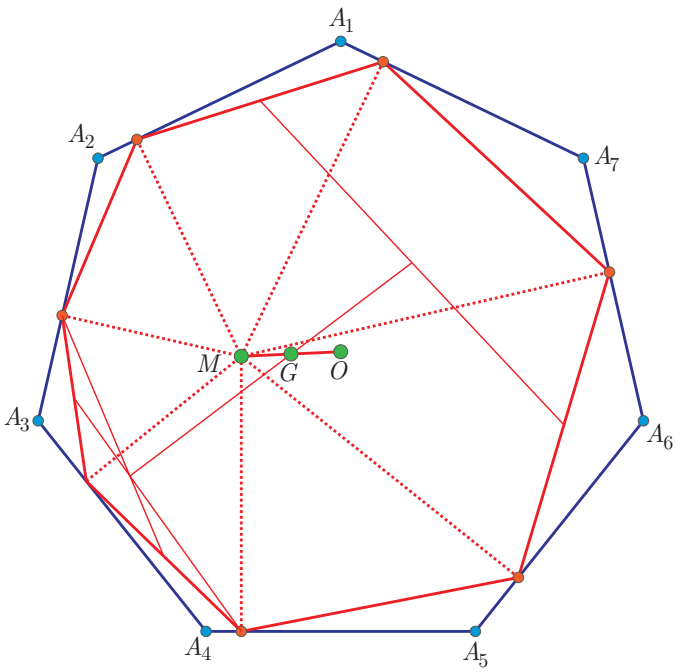


Fig. 2. The centroid of a podar polygon relative to a regular n-gon

Proof. First we consider the case when n is odd. The vertices of the regular n-gon are represented by

$$a_j = \varepsilon^j = \cos\frac{2j\pi}{n} + \sin\frac{2j\pi}{n}, \quad 1 \le j \le n.$$

If $n = 2k+1$, we have $A_kA_{k+1} \parallel OY$ and hence

$$m_k = \cos\frac{2k\pi}{2k+1} + \frac{m-\overline{m}}{2}.$$

Using rotations again we can deduce

$$m_j \cdot \varepsilon^{k-j} = \cos\frac{2k\pi}{2k+1} + \frac{m\cdot\varepsilon^{k-j} - (\overline{m\cdot\varepsilon^{k-j}})}{2}$$

for $1 \le j \le 2k+1$. But $\overline{\varepsilon} = \varepsilon^{2k} = \varepsilon^{-1}$ and hence we have the following relations:

$$m_j = \varepsilon^{k+j+1}\cos\frac{2k\pi}{2k+1} + \frac{m-\varepsilon^{2j+1}\cdot\overline{m}}{2} \quad \text{for } \ 1 \le j \le 2k+1.$$

These relations and the $\sum\limits_{v=0}^{n-1}\varepsilon^v = 0$ equality imply

$$\frac{1}{n}\cdot\sum_{j=1}^{n} m_j = \frac{m}{2}.$$

If n is even ($n = 2k$) we take

$$a_j = z_0\varepsilon^j = z_0\left(\cos\frac{2j\pi}{n} + i\cdot\sin\frac{2j\pi}{n}\right), \quad 1 \le j \le n,$$

where

$$z_0 = \cos\frac{\pi}{2k} + i\cdot\sin\frac{\pi}{2k}$$

and we obtain

$$m_{k-1} = \cos\frac{(2k-1)\pi}{2k} + \frac{m-\overline{m}}{2},$$

$$m_j\cdot\varepsilon^{k-j-1} = \cos\frac{(2k-1)\pi}{2k} + \frac{m\cdot\varepsilon^{k-j-1} - (\overline{m\cdot\varepsilon^{k-j-1}})}{2}$$

for $1 \le j \le 2k$. These relations imply $\frac{1}{n}\cdot\sum\limits_{j=1}^{n} m_j = \frac{m}{2}$ and hence the proof is completed. □

Remark 2.2. *If $n = 2k$, the points corresponding to the complex numbers $\frac{m_j+m_{j+k}}{2}$ are on the axis of symmetry parallel to A_jA_{j+1}. So the relation*

$$\frac{1}{n}\cdot\sum_{j=1}^{n} m_j = \frac{1}{k}\cdot\sum_{j=1}^{k}\frac{m_j+m_{j+k}}{2} = \frac{m}{2}$$

means that the centroid of the k-gon determined by the projections of M to the these axis of symmetry is the midpoint of OM. In this case the property can be viewed as theorem 2.1 with a degenerated regular k-gon (which degenerates into the center of symmetry).

The previous remark suggests that a similar property holds if the points $M_1, M_2, \ldots, M_n$ are the projections of M to the axis of symmetry. To explore this possibility and to formulate a proper conjecture we can construct figures using a geometry software. Finally we obtain the following property:

Theorem 2.3. *The centroid of the n-gon determined by the orthogonal projections of a point M to the axis of symmetry of a regular n-gon is the midpoint of OM (O is the center of symmetry).*

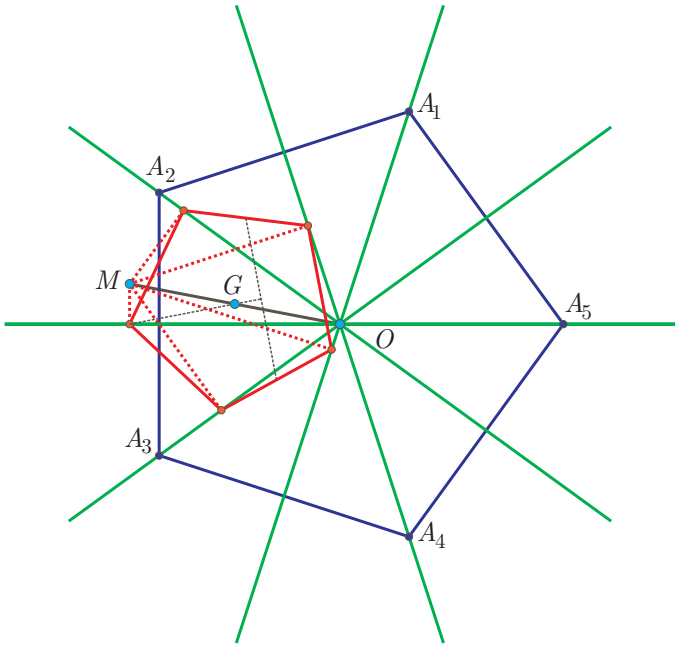


Fig. 3. The centroid of a podar polygon relative to the symmetry axes of a regular n-gon

Proof. The method and the calculations are almost the same. We can assume that one of the symmetry axis is the OY axes, so the real part of the projections affix is 0 while the imaginary part is the same as the imaginary part of the projected point's affix. This implies that if we repeat all the calculations instead of $\cos \frac{(2k-1)\pi}{2k}$ or $\cos \frac{2k\pi}{2k+1}$ we have 0. □

Remark 2.4. *All groups discovered/rediscovered independently theorem 2.1 while theorem 2.3 was observed only by one group when they analyzed their proof.*

In order to generalize the previous results to higher dimensions as a first step we analyze the case of a regular tetrahedron.

Theorem 2.5. *Let $A_1A_2A_3A_4$ be a regular tetrahedron with center O, M_i $(1 \leq i \leq 4)$the orthogonal projections of a point M to it's faces and Q_i $(1 \leq i \leq 6)$ the projections to it's sides. The following two statements are true:*

a) The centroid G_2 of the tetrahedron $M_1M_2M_3M_4$ is on OM and satisfies $\frac{OG_2}{OM} = \frac{2}{3}$.
b) The centroid G_1 of the system $Q_1Q_2Q_3Q_4Q_5Q_6$ is on OM and satisfies $\frac{OG_1}{OM} = \frac{1}{3}$.

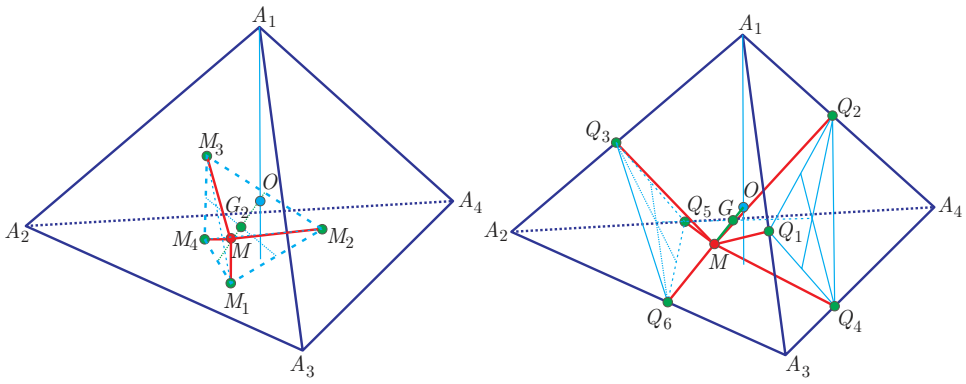


Fig. 4. The centroid of a podar tetrahedron relative to the faces and the sides of a regular tetrahedron

Proof. First we prove that the two statements are equivalent. Due to theorem 1.1. we have (see figure 4):

$$\frac{\overrightarrow{OO_1} + \overrightarrow{OM_1}}{2} = \frac{\overrightarrow{OQ_4} + \overrightarrow{OQ_5} + \overrightarrow{OQ_6}}{3},$$

$$\frac{\overrightarrow{OO_2} + \overrightarrow{OM_2}}{2} = \frac{\overrightarrow{OQ_4} + \overrightarrow{OQ_1} + \overrightarrow{OQ_2}}{3},$$

$$\frac{\overrightarrow{OO_3} + \overrightarrow{OM_3}}{2} = \frac{\overrightarrow{OQ_3} + \overrightarrow{OQ_5} + \overrightarrow{OQ_2}}{3}$$

and

$$\frac{\overrightarrow{OO_4} + \overrightarrow{OM_4}}{2} = \frac{\overrightarrow{OQ_1} + \overrightarrow{OQ_3} + \overrightarrow{OQ_6}}{3},$$

(where $Q_1 \in A_1A_3, Q_2 \in A_1A_4, Q_3 \in A_1A_2, Q_4 \in A_4A_3, Q_5 \in A_2A_4, Q_6 \in A_2A_3$, $\overrightarrow{AB}$ means the vector from A to B, and O_i are the centroids of the faces). From these equalities we deduce

$$\frac{1}{2} \cdot \sum_{i=1}^{4} \overrightarrow{OM_i} = \frac{2}{3} \cdot \sum_{i=1}^{6} \overrightarrow{OQ_i}$$

(because $\frac{1}{2} \cdot \sum_{i=1}^{4} \overrightarrow{OO_i} = 0$) and so

$$\frac{1}{2} \cdot \overrightarrow{OG_2} = \overrightarrow{OG_1}.$$

This relation implies the equivalence of the two statements.

In order to prove the first statement we consider a regular tetrahedron with vertices $A_1\left(0, 0, \frac{\sqrt{6}}{3}\right)$, $A_2\left(-\frac{1}{2}, -\frac{\sqrt{3}}{6}, 0\right)$, $A_3\left(\frac{1}{2}, -\frac{\sqrt{3}}{6}, 0\right)$, $A_4\left(0, \frac{\sqrt{3}}{3}, 0\right)$ and we calculate the coordinates of the points M_i. The equations of the faces are:

$$\begin{aligned}
A_1A_2A_3: &\quad -2\sqrt{3}y + \sqrt{3}z - \sqrt{2} = 0,\\
A_1A_2A_4: &\quad 3\sqrt{2}x - \sqrt{6}y - \sqrt{3}z + \sqrt{2} = 0,\\
A_1A_3A_4: &\quad 3\sqrt{2}x + \sqrt{6}y + \sqrt{3}z - \sqrt{2} = 0 \quad \text{and}\\
A_3A_2A_4: &\quad z = 0
\end{aligned}$$

Using the formula

$$x = x_0 - A \cdot \frac{A \cdot x_0 + B \cdot y_0 + C \cdot z_0 + D}{A^2 + B^2 + C^2},$$

which gives the x coordinate of the orthogonal projection of the point (x_0, y_0, z_0) to the plane with equation $A \cdot x + B \cdot y + C \cdot z + D = 0$ we obtain:

$$x_1 = x_4 = x_0,$$

$$x_2 = x_0 - 3\sqrt{2} \cdot \frac{3\sqrt{2} \cdot x_0 + \sqrt{6} \cdot y_0 + \sqrt{3} \cdot z_0 - \sqrt{2}}{27} \text{ and}$$

$$x_3 = x_0 - 3\sqrt{2} \cdot \frac{3\sqrt{2} \cdot x_0 - \sqrt{6} \cdot y_0 - \sqrt{3} \cdot z_0 + \sqrt{2}}{27}.$$

From these equalities we can easily deduce the relation

$$\frac{x_1 + x_2 + x_3 + x_4}{4} = \frac{2x_0}{3}.$$

A similar calculation shows the same relation for the y and the z coordinates, hence G_2 lies on OP and satisfies the desired equality. This completes the proof. □

Based on this property we can formulate the following conjecture for the n-dimensional euclidian space:

Conjecture 1. *If we denote by G_k the centroid of the system determined by the orthogonal projections of a point M to the k-dimensional faces of a regular n-simplex $A_1A_2A_3...A_nA_{n+1}$, than $G_k \in OM$ and satisfies $\frac{OG_k}{OM} = \frac{k}{n}$, $1 \leq k \leq n-1$, where O is the center of the simplex.*

Remark 2.6. *Theorem 2.5 was discovered/rediscovered by all the groups while conjecture 1 was formulated (without proof) only by one group. After an analysis of the proof of theorem 2.5 all groups formulated the necessity of a different approach in attacking the higher dimensional problem.*

This conjecture can be proved using the same ideas as in the proof of theorem 2.5, but the calculations are more complicated. In order to simplify the proof of this conjecture first we try to give an affine version of the 2 dimensional property and then by extending this affine version to higher dimensions we find a property which is more general than conjecture 1 and admits a simpler proof. For this we need to observe that in a regular polygon (polyhedra or simplex) the segment joining the centroid O of the polygon with the midpoint O_i of a side (the centroid of a face) is perpendicular to this side (face). Hence the construction of the orthogonal projection of M to a side d can be viewed as the construction of a projection in the direction OO_i. This can be done for an arbitrary triangle or even for an arbitrary polygon, so we can formulate the following conjectures:

Conjecture 2. *In the triangle $A_1A_2A_3$ $O_1 \in A_2A_3$, $O_2 \in A_3A_1$ and $O_3 \in A_1A_2$ are the midpoints of the sides and $A_1O_1 \cap A_2O_2 \cap A_3O_3 = \{O\}$. If for an arbitrary point M we consider the points $M_1 \in A_2A_3$, $M_2 \in A_3A_1$ and $M_3 \in A_1A_2$ such that $MM_1||OO_1$, $MM_2||OO_2$ and $MM_3||OO_3$, then the centroid of the triangle $M_1M_2M_3$ is the midpoint of the segment OM.*

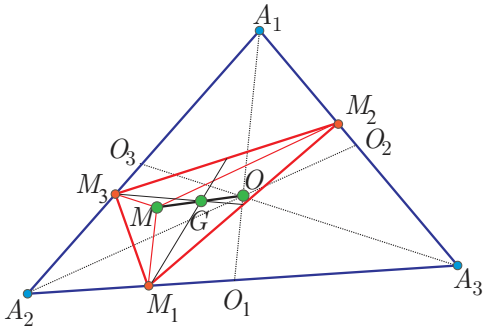


Fig. 5. The centroid of a generalized podar triangle relative to an arbitrary triangle

Conjecture 3. *Denote by $O_1, O_2, \ldots, O_n$ the midpoints of the sides A_1A_2, $A_2A_3, \ldots,$ A_nA_1 in the polygon $A_1A_2 \ldots A_n$ and by O the centroid of the poly-*

gon. If for an arbitrary point M we consider the points $M_i \in A_iA_{i+1}$, $1 \leq i \leq n$ such that $MM_i||OO_i$, $1 \leq i \leq n$, then the centroid of the polygon $M_1M_2 \dots M_n$ is the midpoint of the segment OM.

Using a dynamic geometry software we can explore the validity of these conjectures. Conjecture 2 seems to be true but unfortunately Conjecture 3 is not true for all polygons.

Remark 2.7. *It is an open question to assure conditions for which conjecture 3 becomes a theorem. One of the groups observed that conjecture 3 becomes a theorem for $n = 4$ if $A_1A_2A_3A_4$ is a trapezoid.*

Proof of conjecture 2. If we denote by the corresponding small letter the position vector of each point, we have the following relations: $o = \frac{1}{3}(a_1 + a_2 + a_3)$, $o_1 = \frac{1}{2}(a_2 + a_3)$, $o_2 = \frac{1}{2}(a_3 + a_1)$ and $o_3 = \frac{1}{2}(a_1 + a_2)$. M is an arbitrary point in the plane so there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ such that

$$m = \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3$$

and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Since $M_1 \in A_2A_3$ there exists $\lambda_1 \in \mathbb{R}$ with the property $m_1 = \lambda_1 a_2 + (1 - \lambda_1)a_3$, hence the condition $MM_1||OO_1$ can be expressed as

$$m - m_1 = c(o - o_1)$$

with $c \in \mathbb{R}$. This implies

$$\alpha_1 a_1 + (\alpha_2 - \lambda_1)\, a_2 + (\alpha_3 - 1 + \lambda_1)\, a_3 = c\left(\frac{1}{3}a_1 - \frac{1}{6}a_2 - \frac{1}{6}a_3\right).$$

If the origin of the position vectors is not in the same plane as the vertices of the initial triangle, than a_1, a_2 and a_3 are linearly independent vectors, so we obtain the following system

$$\begin{cases} \alpha_1 & = \frac{1}{3}c \\ \alpha_2 - \lambda_1 & = -\frac{c}{6} \\ \alpha_3 - 1 + \lambda_1 & = -\frac{c}{6} \end{cases}$$

and so

$$m_1 = m - 3\alpha_1(o - o_1) = \left(\alpha_2 + \frac{1}{2}\alpha_1\right) a_2 + \left(\alpha_3 + \frac{1}{2}\alpha_1\right) a_3.$$

Using a similar argument we obtain

$$m_2 = \left(\alpha_3 + \frac{1}{2}\alpha_2\right) a_3 + \left(\alpha_1 + \frac{1}{2}\alpha_2\right) a_1 \text{ and } m_3 = \left(\alpha_1 + \frac{1}{2}\alpha_3\right) a_1 + \left(\alpha_2 + \frac{1}{2}\alpha_3\right) a_2$$

so

$$g = \frac{1}{3}\,(m_1 + m_2 + m_3) = \frac{1}{2}(m + o). \qquad \square$$

Analyzing theorem 2.5 and conjecture 2 we can formulate the following property:

Theorem 2.8. *(Zsolt Szilágyi, Szilárd András) Consider the n-simplex $A_0 \dots A_n$, an arbitrary point M and O the centroid of the simplex. Denote by $O_{i_0 \dots i_k}$ the centroid of the face $A_{i_0} \dots A_{i_k}$ and by $M_{i_0 \dots i_k}$ the intersection of the face $A_{i_0} \dots A_{i_k}$ with the line drawn through M and parallel to $O_{i_0 \dots i_k} A_{i_{k+1} \dots A_{i_n}}$. If $G_k^{'}$ denotes the centroid of the system $M_{i_0 \dots i_k}$, where k is fixed and $i_0 \dots i_k$ takes all possible values, then for each $k \in \{1, 2, 3 \dots, n-1\}$ the points $M, O, G_k^{'}$ are on the same line and*

$$\frac{OG_k^{'}}{OM} = \frac{k}{n}.$$

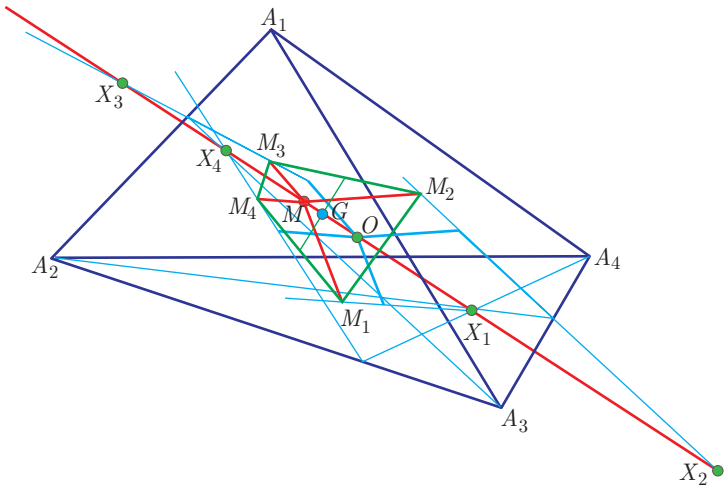


Fig. 6. The centroid of a podar tetrahedron relative to an arbitrary tetrahedron

Proof. Denote by the corresponding small letters the position vectors (in $\mathbb{R}^n$) of the points. M is an arbitrary point in $\mathbb{R}^n$, so there exist $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ such that $\sum_{i=0}^{n} \alpha_i = 1$ and

$$m = \sum_{i=0}^{n} \alpha_i a_i.$$

If $\overline{m} := m_{i_0 \dots i_k} = \sum_{j=0}^{k} c_j a_{i_j}$ (with $\sum_{j=0}^{k} c_j = 1$) is the intersection of the face $A_{i_0} \dots A_{i_k}$ with the parallel line to $O_{i_0 \dots i_k} A_{i_{k+1}} \dots A_{i_n}$, and $o_{i_0 \dots i_k} = \frac{1}{k+1} \sum_{j=0}^{k} a_{i_j}$ is the centroid of the face $A_{i_0} \dots A_{i_k}$, then the condition of parallelism is:

$$m - \overline{m} = c \left(\sum_{j=k+1}^{n} \lambda_j (a_{i_j} - o_{i_0 \dots i_k}) \right),$$

where $\lambda_j \in \mathbb{R}$, for all $j \in \{k+1, \ldots, n\}$ and $\sum\limits_{j=k+1}^{n} \lambda_j = 1$. This can be written as

$$\sum_{j=0}^{k}(\alpha_{i_j} - c_j)a_{i_j} + \sum_{j=k+1}^{n} \alpha_{i_j} a_{i_j} = c\left(\sum_{j=k+1}^{n} \lambda_j a_{i_j} - \frac{1}{k+1}\sum_{j=0}^{k} a_{i_j}\right),$$

so we deduce $\lambda_j = \frac{\alpha_{i_j}}{c}$, for $k+1 \leq j \leq n$, and $c_j = \alpha_{i_j} + \frac{c}{k+1}$, for $0 \leq j \leq k$. From these relations we obtain $c = \sum\limits_{j=k+1}^{n} \alpha_{i_j}$, and so

$$\overline{m} = m_{i_0 \ldots i_k} = \sum_{j=0}^{k}\left[\alpha_{i_j} + \frac{1}{k+1}(\alpha_{i_{k+1}} + \ldots + \alpha_{i_n})\right] a_{i_j}.$$

The number of the k dimensional faces in an n-simplex is C_{n+1}^{k+1}, hence

$$\begin{aligned} g_k^{'} &= \frac{1}{C_{n+1}^{k+1}} \sum_{i \in C_{n+1,k+1}} m_{i_0 \ldots i_k} \\ &= \frac{1}{C_{n+1}^{k+1}} \sum_{i \in C_{n+1,k+1}} \left(\sum_{j=0}^{k}\left[\alpha_{i_j} + \frac{1}{k+1}(\alpha_{i_{k+1}} + \ldots + \alpha_{i_n})\right] a_{i_j}\right), \end{aligned}$$

where $C_{n+1,k+1}$ denotes the set of all combinations of order $(k+1)$ formed from the set $\{0, 1, \ldots, n+1\}$. The number of combinations i for which $l \in \{i_0, \ldots, i_k\}$, where l is a fixed index, is C_n^k while the number of combinations i for which $l \in \{i_0, \ldots, i_k\}$ and $j \in \{i_{k+1}, \ldots, i_n\}$ is C_{n-1}^k. This implies

$$\begin{aligned} g_k^{'} &= \frac{1}{C_{n+1}^{k+1}} \sum_{l=0}^{n}\left[C_n^k \alpha_l + \sum_{j \neq l} \frac{C_{n-1}^k}{k+1}\alpha_j\right] a_l \\ &= \frac{1}{C_{n+1}^{k+1}} \sum_{l=0}^{n}\left[C_n^k \alpha_l + \frac{C_{n-1}^k}{k+1}(1-\alpha_l)\right] a_l \\ &= \sum_{l=0}^{n}\left(\frac{k+1}{n+1}\alpha_l + \frac{n-k}{n(n+1)}(1-\alpha_l)\right) a_l \\ &= \sum_{l=0}^{n}\left(\frac{k}{n}\alpha_l + \frac{n-k}{n(n+1)}\right) a_l. \end{aligned}$$

The centroid is $o = \frac{1}{n+1}\sum\limits_{i=0}^{n} a_i$, so

$$\begin{aligned} \overrightarrow{OG_k^{'}} &= \sum_{l=0}^{n}\left(-\frac{1}{n+1} + \frac{k}{n}\alpha_l + \frac{n-k}{n(n+1)}\right) a_l \\ &= \sum_{l=0}^{n}\left(-\frac{k}{n(n+1)} + \frac{k}{n}\alpha_l\right) a_l. \end{aligned} \tag{1}$$

But $\overrightarrow{OM} = \sum_{l=0}^{n} \left(\alpha_l - \frac{1}{n+1}\right) a_l$, hence $\overrightarrow{OG'_k} = \frac{k}{n}\overrightarrow{OM}$, which completes the proof. □

Remark 2.9. *Theorem 2.8 proves Conjecture 1, because in a regular simplex $O_{i_0\dots i_k}A_{i_{k+1}}\dots A_{i_n}$ is orthogonal to the face $A_{i_0}\dots A_{i_k}$. At the training course this theorem and it's proof was presented by the instructors.*

Examining theorem 2 the following natural question arise: can we replace the midpoints of the faces by other points? The following theorem answers this question.

Theorem 2.10. *Consider a triangle $A_1A_2A_3$ and the real numbers w_1, w_2, w_3 with $w_1 + w_2 + w_3 = 1$. Let O be the point with barycentric coordinates (w_1, w_2, w_3) and M an arbitrary point in the plane of the triangle. If we construct the points $M_1 \in A_2A_3$, $M_2 \in A_3A_1$ and $M_3 \in A_1A_2$ such that $MM_1||OA_1$, $MM_2||OA_3$ and $MM_3||OA_3$, then the centroid of the triangle $M_1M_2M_3$ coincides with the centroid of the triangle MOP, where the barycentric coordinates of P relative to $M_1M_2M_3$ are (w_1, w_2, w_3).*

Remark 2.11. *If $w_1 = w_2 = w_3$, then P is the centroid of $M_1M_2M_3$, so the above theorem reduces to conjecture 2.*

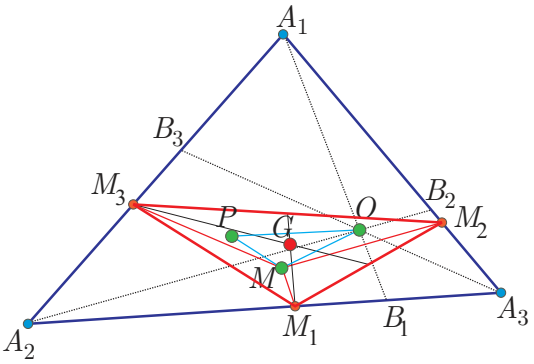


Fig. 7. The centroid of a generalized podar triangle relative to an arbitrary triangle

Proof. If $\gamma_1, \gamma_2, \gamma_3$ denote the barycentric coordinates of M, then

$$m = \gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3$$

and

$$o = w_1 a_1 + w_2 a_2 + w_3 a_3,$$

where a_1, a_2, a_3 are the position vectors of the vertices and m, o the corresponding position vectors of the points M and O. $M_1 \in A_2A_3$, so there exists $\lambda_1 \in \mathbb{R}$

such that $m_1 = \lambda_1 a_2 + (1-\lambda_1)a_3$. The condition $MM_1 || OA_1$ can be expressed as $m_1 - m = c(o - a_1)$, with some $c \in \mathbb{R}$. This implies

$$\lambda_1 a_2 + (1-\lambda_1)a_3 - \gamma_1 a_1 - \gamma_2 a_2 - \gamma_3 a_3 = c(w_1 a_1 + w_2 a_2 + w_3 a_3 - a_1) \quad (2)$$

If we suppose that the starting point of the position vectors is outside the plane $A_1A_2A_3$, then a_1, a_2, a_3 are linearly independent, hence (2) implies $-\gamma_1 = c(w_1 - 1)$, $\lambda_1 - \gamma_2 = cw_2$ and $1 - \lambda_1 - \gamma_3 = cw_3$. From these relations we obtain

$$m_1 = \left(\gamma_2 + \gamma_1 \frac{w_2}{w_2 + w_3}\right) a_2 + \left(\gamma_3 + \gamma_1 \frac{w_3}{w_2 + w_3}\right) a_3. \quad (3)$$

By a similar reasoning we deduce

$$m_2 = \left(\gamma_3 + \gamma_2 \frac{w_3}{w_1 + w_3}\right) a_3 + \left(\gamma_1 + \gamma_2 \frac{w_1}{w_1 + w_3}\right) a_1, \quad (4)$$

$$m_3 = \left(\gamma_1 + \gamma_3 \frac{w_1}{w_2 + w_1}\right) a_1 + \left(\gamma_2 + \gamma_3 \frac{w_2}{w_2 + w_1}\right) a_2. \quad (5)$$

From (3), (4) and (5) we deduce

$$m_1(1 - w_1) + m_2(1 - w_2) + m_3(1 - w_3) = m + o,$$

thus

$$\frac{1}{3}(m_1 + m_2 + m_3) = \frac{1}{3}(m + o + p),$$

which is the desired property. □

Remark 2.12. *Theorem 2.10 can be extended also to higher dimensional simpleces.*

Remark 2.13. *It would be interesting to find an affine version which generalizes all the previous properties (including theorem 2.1 too).*

3. Concluding remarks

- From scientific point of view the training course was fruitful because the participants discovered new properties of the studied configurations and they developed new generalizations, which clarify some aspects of the problem.
- Remark 2.7 and remark 2.13 shows that interesting questions arose, they can constitute a starting point for further study.
- From a teaching point of view the main aim of the training course was achieved, the participants got familiar with the inquiry based methods, they realized that a deeper inquiry of the problems can constitute a solid motivational base for further individual study. They also perceived that in many cases the understanding of the background implicitly requires a deeper inquiry and this can be done even in the traditional framework and context.

- Some participants pointed out that the major inference of this training course was that they realized the importance of simultaneous alternatives in solving a problem and the importance of selecting the tools they use in designing the proofs of well known properties.
- The participants emphasized that a major task in preparing and designing an inquiry based lesson is the selection of a sufficiently rich problem.

Acknowledgements

This paper is based on the work within the project Primas. Coordination: University of Education, Freiburg. Partners: University of Genève, Freudenthal Institute, University of Nottingham, University of Jaen, Konstantin the Philosopher University in Nitra, University of Szedged, Cyprus University of Technology, University of Malta, Roskilde University, University of Manchester, Babeş-Bolyai University, Sør-Trøndelag University College. The author wishes to thank his students and colleagues attending the training course organized by the Babeş-Bolyai University in the framework of the FP7 project PRIMAS[a]. The author was partially supported by the Hungarian University Federation of Cluj Napoca.

References

András, Sz. (2001): About the centroid of podar polygons, Octogon Mathematical Magazine, 9:2, 864.-868.

András, Sz., Szilágyi, Zs. (2006): Geometria II (in Hungarian), Státus, Csíkszereda

Andreescu, T., Andrica, D. (2005): Complex Numbers from A to ...Z, Birkhuser Boston

Dincă, M., Chiriţă M. (1996): Aplicaţii ale numerelor complexe în geometrie (in Romanian), Editura All Educaţional

[a]Promoting inquiry in Mathematics and science education across Europe, Grant Agreement No. 244380

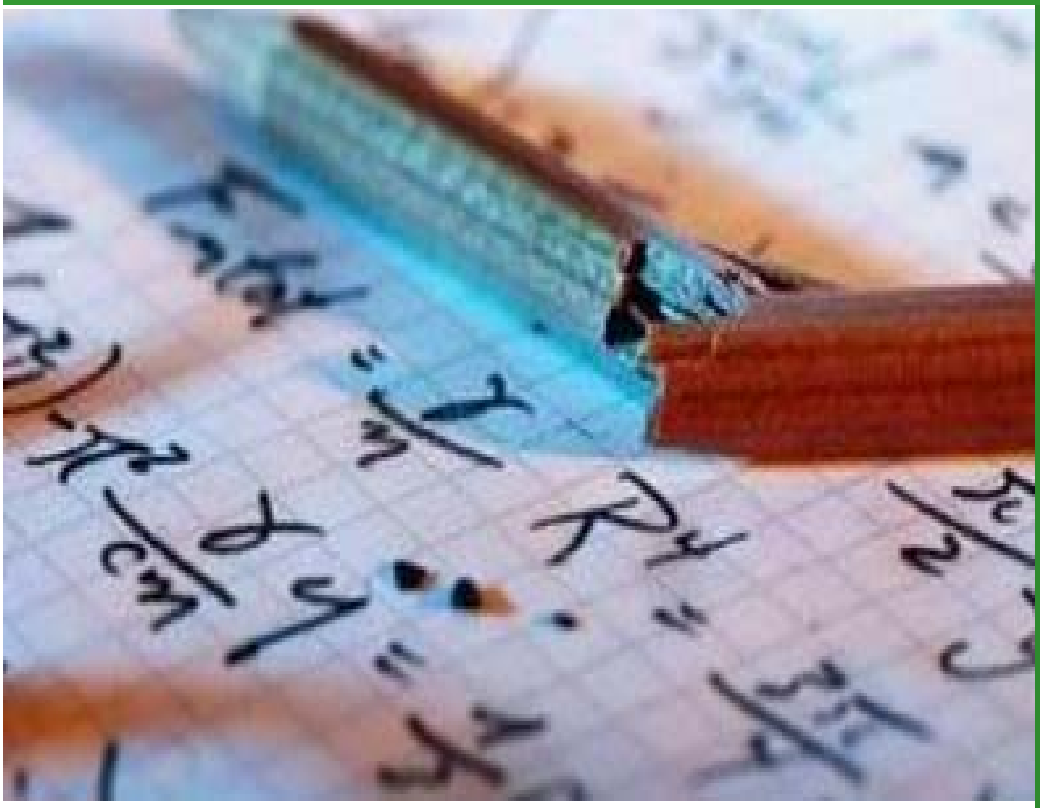
ISSN 0974 – 3340

SUTRA

International Journal of

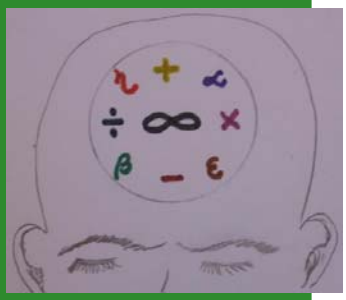
Mathematical Science Education

VOLUME 4, ISSUE II JULY-DECEMBER 2011



Published by

Technomathematics Research Foundation, India



EDITORIAL BOARD

Editorial Advisor:

Arvindkumar, *Professor & Director, Homi Bhabha Centre for Science Education, TIFR, Mumbai, India.*

Editorial Coordinator:

Rupali Akerkar *Technomathematics Research Foundation, Kolhapur, India.*

Editors:

Carmen Batanero, *University of Granada, Campus de Cartuja, Granada, Spain.*
Lokenath Debnath, *Chair Professor of Mathematics, University of Texas-Pan American, Edinburg, Texas, USA.*
Paul Ernest, *Emeritus Professor of Mathematics Education, University of Exeter, UK.*
Laxman Hegde, *Professor of Statistics, Frostburg State University, Frostburg, USA.*
Paulus Gerdes, *Vice-President for S. Africa, African Academy of Sci., Research Centre for Mathematics, Culture & Education, Mozambique.*
A. Krishnamoorthy, *Professor & Head, Mathematics Department, Cochin University of Science & Technology, Kochi, India.*
Hee-chan Lew, *Professor of Mathematics Education, Korea National University of Education, Chungbuk, South Korea.*
Victor Ohanyan, *Professor of Mathematics, American University of Armenia, Armenia.*
D. P. Patil, *Professor of Mathematics, Indian Institute of Science, Bangalore, India.*
L. Radhakrishna, *Retd. Professor & Head, Department of Mathematics Shivaji University, India.*
Geetha S.Rao, *Retd. Professor, Ramanujan Institute, University of Madras, Chennai, India.*
V Ravichandran, *Reader, Department of Mathematics, University of Delhi, India*
Fernando Díez Rubio, *Departamento de Ingeniería Informática, Universidad Autonoma de Madrid, Spain.*
Anand P. Singh, *Professor of Mathematics, University of Jammu, Jammu, India.*
Arindama Singh, *Professor of Mathematics, Indian Institute of Technology Madras, Chennai, India.*
Nada Stehlikova, *Faculty of Education, Charles University, Prague, Czech Republic.*
K. G. Subramanian, *Professor of Mathematics, Universiti Sains Malaysia, Malaysia.*
Mike Thomas, *A/Prof. & Head, Mathematics Education Unit, The University of Auckland, Auckland, New Zealand.*

Editorial Office: Technomathematics Research Foundation,
204/17 Kh, New Shahupuri, Kolhapur 416001, India.
E-mail: director@tmrfindia.org

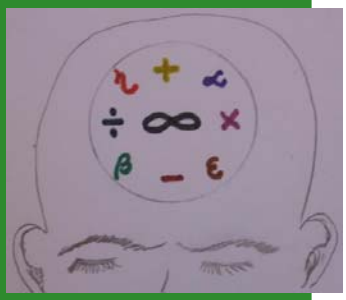
Copyright ©2011 ***Technomthematics Research Foundation.***

All rights reserved. This journal issue or parts thereof may not be reproduced in any form or by any means, electrical or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the copyright owner.

Permission to quote from this journal is granted provided that the customary acknowledgement is given to the source.

Sutra: International Journal of Mathematical Science Education (ISSN 0974 – 3340) is high quality electronic journal published six-monthly by Technomathematics Research Foundation, Kolhapur, India.
URL: www.tmrfindia.org/sutra.html.

Sutra: International Journal of Mathematical Science Education



CONTENTS (Vol. 4, No. 2, 2011)

1 **CHORD LENGTH DISTRIBUTION FUNCTION FOR CONVEX POLYGONS** 1—15
H. S. HARUTYUNYAN AND V. K. OHANYAN

2 **TEACHIING ADAPTIVE AND STRATEGIC REASONING THROUGH FORMULA DERIVATION: BEYOND FORMAL SEMIOTICS** 16—26
ELLIOTT OSTLER

3 **SOLVING CUBICS BY DIVISION METHOD** 27—29
RAGHAVENDRA G. KULKARNI

4 **CORRECTION TO "A NEW METHOD FOR SOLVING QUARTICS"** 30—31
KENNETH V. CARTWRIGHT

Sutra: International Journal of Mathematical Science Education
© Technomathematics Research Foundation
Vol. 4 No. 2, pp. 1 - 15, 2011

CHORD LENGTH DISTRIBUTION FUNCTION FOR CONVEX POLYGONS

H. S. HARUTYUNYAN
Department of Mathematics and Mechanics,
Yerevan State University, Alex Manoogian, 1
0025 Yerevan, Armenia.
hrach87@web.am.

V. K. OHANYAN*
Department of Mathematics and Mechanics,
Yerevan State University, Alex Manoogian, 1
0025 Yerevan, Armenia.
victo@aua.am.

Using the inclusion-exclusion principle and the Pleijel identity, an algorithm for calculation chord length distribution function for a bounded convex polygon is obtained. In the particular case an expression for the chord length distribution function for a rhombus is obtained.

Keywords: chord length distribution function; Pleijel identity; rhombus.

1. INTRODUCTION

Let G be the space of lines g in the Euclidean plane R^2, (p, φ) = the polar coordinates of the foot of the perpendicular to g from the origin O, be standard coordinates for a line $g \in \mathrm{G}$.

Let $\mu(\cdot)$ stand for locally finite measure on G invariant with respect to the group of all Euclidean motions (translations and rotations). It is well known that the element of the measure up to a constant factor has the following form (see [1], [2]):

$$\mu(dg) = dg = dp\, d\varphi,$$

where dp is one dimensional Lebesgue measure, while $d\varphi$ is the uniform measure on the unit circle.

For each bounded convex domain D the set of lines that intersect D we denote by

$$[\mathrm{D}] = \{g \in \mathrm{G} : \ g \cap D \neq \emptyset\}$$

*This research was partially supported by the State Committee of Science of the Republic of Armenia, Grant 11-1A-125.

and we have (see [1], [2]):

$$\mu([\mathbb{D}]) = |\partial\mathbb{D}|,$$

where $\partial\mathbb{D}$ is the boundary of $\mathbb{D}$ and $|\partial\mathbb{D}|$ stands for the length of $\partial\mathbb{D}$.
Let $A^y_{\mathbb{D}}$ be the set of lines that intersect $\mathbb{D}$ producing a chord $\chi(g) = g \bigcap \mathbb{D}$ of length less than or equal to y:

$$A^y_{\mathbb{D}} = \{g \in [\mathbb{D}] : \ |\chi(g)| \leq y\}, \qquad y \in \mathbb{R}.$$

Distribution function of the length of a random chord χ of $\mathbb{D}$ is defined as

$$F(y) = \frac{1}{|\partial\mathbb{D}|}\,\mu(A^y_{\mathbb{D}}) = \frac{1}{|\partial\mathbb{D}|}\iint_{A^y_{\mathbb{D}}} d\varphi\, dp. \tag{1.1}$$

Therefore, to obtain chord length distribution function for a bounded convex domain $\mathbb{D}$ we have to calculate the integral in the right-hand side of (1.1). Explicit formulae for the chord length distribution functions are known only for the cases of a disc, a rectangle [3], and a regular polygon [5].
The main result of the paper is an algorithm for calculation the chord length distribution function for a bounded convex polygon. In particular, an expression for the chord length distribution function for a rhombus is obtained.
The determination of the chord length distribution function has a long tradition of application to collections of bounded convex bodies forming structures in metal and ceramics. The series of formulae for chord length distribution functions may be of use in finding suitable models when empirical distribution functions are given (see [5], [6], [9] and [12]).
In [11] Sulanke proved that $\chi(g) = g \cap D$ is a measurable function on $[D]$. He also proved that the function $F(y)$ defined by formula (1.1) is a continuous function with respect to y. In [8] Gates obtained some properties and inequalities for the chord length distribution function for any bounded convex domain on the plane. In particular, he proved that in the case of a bounded convex polygon the chord length distribution function is expressed in terms of elementary functions.

2. THE CASE OF A CONVEX POLYGON

Let $\mathbb{D}$ be a convex bounded polygon in the plane and a_1, a_2,..., a_n be sides of $\mathbb{D}$. Then

$$[\mathbb{D}] = \bigcup_{i<j}([a_i] \cap [a_j])$$

where $[a_i] \cap [a_j]$ is the set of lines hitting both sides a_i and a_j of $\mathbb{D}$.
We write

$$F(y) = \frac{1}{|\partial\mathbb{D}|}\sum_{i<j}\iint_{\{g\in[a_i]\cap[a_j]:|\chi(g)|\leq y\}} d\varphi\, dp =$$

$$= \frac{1}{|\partial\mathbb{D}|}\left[\sum_{i<j}^{I}\iint_{\{g\in[a_i]\cap[a_j]:|\chi(g)|\le y\}} d\varphi\,dp + \sum_{i<j}^{II}\iint_{\{g\in[a_i]\cap[a_j]:|\chi(g)|\le y\}} d\varphi\,dp\right],$$

where $\sum^{I}$ is over all pairs of nonparallel segments a_i, $a_j \subset \partial\mathbb{D}$ and $\sum^{II}$ is over all pairs of parallel segments a_i and $a_j \subset \partial\mathbb{D}$.
In each of the integrals in the sum $\sum_{i<j}^{I}$ one integration can be performed by passing to the $(|\chi|, \varphi)$ coordinates. We have (see [2], page 157):

$$dg = \frac{\sin\alpha_1\,\sin\alpha_2}{|\sin(\alpha_1+\alpha_2)|}\,d|\chi|\,d\varphi, \tag{2.1}$$

where α_1 is the angle between a_i and $\chi(g) = g\cap\mathbb{D}$, while α_2 is the angle between a_j and $\chi(g)$ ($g \in [a_i]\cap[a_j]$), α_1 and α_2 lie in one half-plane with respect to inside of $\mathbb{D}$.

3. PLEIJEL IDENTITY

Let $\mathbb{D}$ be a convex bounded polygon in the plane and a_1,..., a_n be sides of $\mathbb{D}$. The so–called Pleijel identity for $\mathbb{D}$ is as follows (see [2], page 156):

$$\int_{[\mathbb{D}]} f(|\chi(g)|)\,dg = \int_{\mathbb{G}} f'(|\chi|)\cdot|\chi|\,\cot\alpha_1\,\cot\alpha_2\,dg + \sum_{i=1}^{n}\int_0^{|a_i|} f(u)\,du, \tag{3.1}$$

where $f(x)$ is a function with continuous first derivative $f'(x)$, α_1 and α_2 are the angles between $\partial\mathbb{D}$ and g at the endpoints of $\chi(g) = g\cap\mathbb{D}$ which lie in one half-plane with respect to the inside of $\mathbb{D}$, $|a_i|$ is the length of a_i, $i = 1, ..., n$.

It was shown by R. V. Ambartzumian in [2] (page 156) that the identity (3.1) is useful to calculate chord length distribution function for the case of a bounded convex polygon. If in (3.1) we formally put

$$f_y(u) = \begin{cases} 0 & \text{if} \quad u \le y \\ 1 & \text{if} \quad u > y \end{cases}$$

then the left-hand integral in (3.1) will equal

$$\mu\{g \in [\mathbb{D}] :\ |\chi(g)| > y\},$$

i.e. the invariant measure of the set of chords of $\mathbb{D}$ whose length exceeds y. The derivative of $f_y(u)$ should be replaced by Dirac's δ–function concentrated at y (see [2], page 156). Therefore, we obtain

$$[1-F(y)]\,|\partial\mathbb{D}| = \sum_{i<j}^{I}\iint_{[a_i]\cap[a_j]}\delta(|\chi|-y)\cdot|\chi|\,\cot\alpha_1\,\cot\alpha_2\,dg +$$

$$+\sum_{i<j}^{II}\iint_{[a_i]\cap[a_j]}\delta(|\chi|-y)\cdot|\chi|\,\cot\alpha_1\,\cot\alpha_2\,dg + \sum_{i=1}^{n}(|a_i|-y)^+, \tag{3.2}$$

where $x^+ = x$ if $x > 0$, and 0 otherwise.

For any continuous function f we have (see [10]):

$$\int_{\mathbb{R}^n} \delta(x-y)\, f(x)\, dx = f(y). \tag{3.3}$$

For each pair $\{a_i, a_j\} \in \sum_{i<j}^{I}$ we make the change of variables $(p,\varphi) \to (|\chi|,\varphi)$ and using (2.1) and (3.3) we obtain

$$\iint_{[a_i]\cap[a_j]} \delta(|\chi|-y)\cdot|\chi|\,\cot\alpha_1\,\cot\alpha_2\, dg = \frac{y}{\sin\gamma_{ij}} \int_{\Phi_{ij}(y)} \sin\varphi\,\sin(\gamma_{ij}-\varphi)\, d\varphi,$$

where γ_{ij} is the angle between nonparallel sides a_i and a_j (or their continuations), φ is the angle between direction φ and direction of segment a_i $i<j$, and

$$\Phi_{ij}(y) = \{\varphi : \text{a chord joining}\, a_i \,\text{and}\, a_j \,\text{exists with direction}\, \varphi \,\text{and length}\, y\}.$$

Note, that $\Phi_{ij}(y)$ is a subset in the space of directions in the plane, while in the integral in the right hand side the corresponding $\Phi_{ij}(y)$ is the set of angles, where reference direction coincides with direction of segment a_i, $i<j$ and reference origin coincides with the intersection of the lines containing a_i and a_j.

Moreover, for parallel sides a_i and a_j (i.e. a_i, $a_j \in \sum_{i<j}^{II}$) we have

$$\sum_{i<j}^{II} \iint_{[a_i]\cap[a_j]} \delta(|\chi|-y)\cdot|\chi|\,\cot\alpha_1\,\cot\alpha_2\, dg =$$

$$= -\sum_{i<j}^{II} I_{ij}(y) \int \delta(|\chi|-y)\cdot|\chi|\, h(\varphi_\chi)\,\tan^2\varphi_\chi\, \frac{d\varphi_\chi}{d|\chi|}\, d|\chi| =$$

$$= -\sum_{i<j}^{II} I_{ij}(y)\, h(\varphi_y)\,\tan^2\varphi_y\, \frac{b_{ij}}{\sqrt{y^2-b_{ij}^2}},$$

where φ_y takes on two values $\arccos\frac{b_{ij}}{y}$ or $2\,\pi - \arccos\frac{b_{ij}}{y}$, b_{ij} is the distance between the parallel segments a_i and a_j (i.e. the distance between the lines containing a_i and a_j), and $I(A)$ is the indicator function of the event A, i.e. $I(A)=1$ if A has occurred and 0 otherwise, indicator $I_{ij}(y) = I(\text{length of shortest chord hitting}\, a_i \,\text{and}\, a_j \le y \le \text{length of longest chord hitting}\, a_i \;\text{and}\, a_j)$. Further, $h(\varphi) \neq b_{ij}$ is the height of the maximal parallelogram with two sides equal to $\chi(\varphi) = g(\varphi)\cap\mathbb{D}$, $g(\varphi)\in[a_i]\cap[a_j]$ ($g(\varphi)$ is a line with φ-direction), and the other two sides lie on the parallel sides a_i and a_j,

$$h(\varphi_\chi) = h\left(\arccos\frac{b_{ij}}{|\chi(\varphi)|}\right) + h\left(2\,\pi - \arccos\frac{b_{ij}}{|\chi(\varphi)|}\right).$$

Hence, $h(\cdot)=0$ if the parallelogram is empty.
For the value of φ such that $|\chi(\varphi)| = y$ we have $h(\varphi_y) = h(\varphi_\chi)$.

Therefore we obtain

$$F(y) = 1 - \frac{1}{\sum\limits_{i=1}^{n} |a_i|} \left[\sum_{i<j}^{I} \frac{y}{\sin\gamma_{ij}} \int_{\Phi_{ij}(y)} \sin\varphi \, \sin(\gamma_{ij} - \varphi)\, d\varphi - \right.$$
$$\left. - \sum_{i<j}^{II} I_{ij}(y)\, h(\varphi_y)\, \frac{\sqrt{y^2 - b_{ij}^2}}{b_{ij}} + \sum_{i=1}^{n} (|a_i| - y)^+ \right]. \qquad (3.4)$$

In the case where $\partial\mathbb{D}$ contains no pairs of parallel sides, formula (3.4) coincides with the expression given by R. V. Ambartzumian in [2], page 158.

It follows from (3.4) that to find distribution function $F(y)$ we have to calculate integrals of the form

$$\frac{1}{\sin\gamma} \int_{\Phi_{a,b}(y)} \sin\varphi \, \sin(\gamma - \varphi)\, d\varphi$$

for any two nonparallel segments a and b ($b \le a$) with the angle γ between a and b (or their continuations) and also calculate the second sum in (3.4) (for pairs of parallel sides). Here and below $\Phi_{a,b}(y)$ is

$$\Phi_{a,b}(y) = \{\varphi : \text{a chord joining } a \text{ and } b \text{ exists with direction } \varphi \text{ and length } y\}.$$

4. THE DOMAIN $\Phi_{a,b}(y)$

Consider segments a and b ($b \le a$), which have one common endpoint and make an angle γ.

The line g with parameters (φ, p) intersect sides a and b, if

$$\varphi \in \left(-\frac{\pi}{2} + \gamma, \frac{\pi}{2}\right).$$

Without loss of generality we can assume, that the reference direction (X-axis) coincides with the direction of the segment a and the origin O coincides with the intersection of the lines containing a and b. Find the intersection point of the line $g = (\varphi, p)$ with the segment a. We have

$$x = \frac{p}{\cos\varphi} \quad \text{and} \quad y = 0.$$

Therefore,

$$g \cap a = \begin{cases} x = \frac{p}{\cos\varphi} \\ y = 0 \\ 0 \le p \le a\, \cos\varphi. \end{cases}$$

Find the intersection point of the line $g = (\varphi, p)$ with the segment b. We have

$$\begin{cases} x\, \cos\varphi + y \sin\varphi = p \\ y = \tan\gamma\, x \end{cases} \qquad \text{or} \qquad \begin{cases} x = \frac{p\, \cos\gamma_k}{\cos(\varphi - \gamma)} \\ y = \frac{p\, \sin\gamma}{\cos(\varphi - \gamma)} \end{cases} .$$

Hence, we obtain

$$g \cap b = \begin{cases} x = \frac{p\,\cos\gamma}{\cos(\varphi-\gamma)} \\ y = \frac{p\,\sin\gamma}{\cos(\varphi-\gamma)} \\ 0 \le p \le b\cos(\varphi-\gamma). \end{cases}$$

The line g which intersects sides a and b forms a chord χ of the length

$$|\chi| = \frac{p\,\sin\gamma}{\cos\varphi\,\cos(\varphi-\gamma)}.$$

Therefore, we come to the following system of relations:

$$\begin{cases} 0 \le p \le a\cos\varphi \\ 0 \le p \le b\cos(\varphi-\gamma) \\ |\chi| = \frac{p\,\sin\gamma}{\cos\varphi\,\cos(\varphi-\gamma)}. \end{cases} \tag{4.1}$$

If $a\cos\varphi \le b\cos(\varphi-\gamma)$, then $\tan\varphi \ge \frac{a-b\cos\gamma}{b\sin\gamma}$. Hence, $\varphi \ge \arctan\frac{a-b\cos\gamma}{b\sin\gamma}$. It is not difficult to verify, that $\arctan\frac{a-b\cos\gamma}{b\sin\gamma} \in \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right)$. Since $\varphi \in \left(-\frac{\pi}{2}+\gamma, \arctan\frac{a-b\cos\gamma}{b\sin\gamma}\right)$, then (4.1) has the following form:

$$\begin{cases} 0 \le p \le b\cos(\varphi-\gamma) \\ |\chi| = \frac{p\,\sin\gamma}{\cos\varphi\,\cos(\varphi-\gamma)}. \end{cases}$$

Moreover, we get

$$\chi_{\min}(\varphi) = 0, \qquad \chi_{\max}(\varphi) = \frac{b\sin\gamma}{\cos\varphi}.$$

The function $\chi_{\max}(\varphi)$ in the domain $\left(-\frac{\pi}{2}+\gamma, 0\right)$ decreasing, while in the domain $\left(0, \arctan\frac{a-b\cos\gamma}{b\sin\gamma}\right)$ increasing. Hence, we have

$$\varphi_{\min} = 0, \quad \chi_{\max}\left(-\frac{\pi}{2}+\gamma\right) = b, \quad \chi_{\max}(\varphi_{\min}) = b\sin\gamma \quad \text{and}$$

$$\chi_{\max}\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma}\right) = \sqrt{a^2+b^2-2ab\cos\gamma} = c,$$

where c is the third side of the triangle made by a and b.
Now let $\varphi \in (\arctan\frac{a-b\cos\gamma}{b\sin\gamma}, \frac{\pi}{2})$. In this case (4.1) can be rewritten in the form:

$$\begin{cases} 0 \le p \le a\cos(\varphi)) \\ |\chi| = \frac{p\,\sin\gamma}{\cos\varphi\,\cos(\varphi-\gamma)}. \end{cases}$$

Moreover, we get

$$\chi_{\min}(\varphi) = 0, \qquad \chi_{\max}(\varphi) = \frac{a\sin\gamma}{\cos(\varphi-\gamma)}.$$

If $\arctan\frac{a-b\cos\gamma}{b\sin\gamma} > \gamma$ $\left(\text{i.e. } \cos\gamma > \frac{b}{a} \text{ or } \gamma < \arccos\frac{b}{a}\right)$, than the function $\chi_{\max}(\varphi)$ in the domain $\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma}, \frac{\pi}{2}\right)$ increasing,

$$\chi_{\max}\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma} - \gamma\right) = c, \qquad \chi_{\max}\left(\frac{\pi}{2}\right) = a.$$

If $\arctan\frac{a-b\cos\gamma}{b\sin\gamma} \le \gamma$ $\left(\text{i.e. } \cos\gamma \le \frac{b}{a} \text{ or } \gamma \ge \arccos\frac{b}{a}\right)$, than the function $\chi_{\max}(\varphi)$ in the domain $\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma}, \gamma\right)$ decreasing, while in the domain $\left(\gamma, \frac{\pi}{2}\right)$ increasing,

$$\varphi_{\min} = \gamma, \quad \chi_{\max}(\varphi_{\min}) = a\sin\gamma, \quad \chi_{\max}\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma} - \gamma\right) = c, \quad \text{and}$$

$$\chi_{\max}\left(\frac{\pi}{2}\right) = a.$$

Let $\gamma < \arccos\frac{b}{a}$. It is not difficult to verify that $c < b$, if $\cos\gamma > \frac{a}{2b}$, or $\gamma < \arccos\frac{a}{2b}$. From the other hand, $\arccos\frac{a}{2b} < \arccos\frac{b}{a}$, if $a > b\sqrt{2}$.

So, if $a > b\sqrt{2}$ and $\gamma \le \arccos\frac{a}{2b}$, than

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup\left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, c) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup\left(\gamma+\arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [c, b) \\ \left(\gamma+\arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b, a] \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases} \tag{4.2}$$

If $a > b\sqrt{2}$ and $\arccos\frac{a}{2b} < \gamma \le \arccos\frac{b}{a}$, we have

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup\left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, b) \\ \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b, c) \\ \left(\gamma+\arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [c, a] \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases} \tag{4.3}$$

When $a \le b\sqrt{2}$, for every $\gamma < \arccos\frac{b}{a}$, $c < b$, and $\Phi_{a,b}(y)$ is defined by (4.2).

Now let $\arccos\frac{b}{a} \le \gamma < \frac{\pi}{2}$. It follows from the Sine theorem for the triangles, that $c \ge a\sin\gamma$. Moreover, if $a > b\sqrt{2}$, $\arcsin\frac{b}{a} < \arccos\frac{b}{a}$, so for every $\gamma \ge \arccos\frac{b}{a}$, $a\sin\gamma \ge b$.

Therefore, in the case $a > b\sqrt{2}$ and $\arccos\frac{b}{a} \le \gamma < \arccos\frac{b}{2a}$ we have

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, b) \\ \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b, a\sin\gamma) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [a\sin\gamma, c) \\ \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [c, a) \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases} \tag{4.4}$$

When $a > b\sqrt{2}$ and $\arccos\frac{b}{2a} \le \gamma < \frac{\pi}{2}$ we have

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, b) \\ \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b, a\sin\gamma) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [a\sin\gamma, a) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right), & \text{if} \quad y \in [a, c) \\ \emptyset, & \text{if} \quad y > c \quad \text{or} \quad y < 0 \end{cases} \tag{4.5}$$

Let $a \le b\sqrt{2}$ and $\arccos\frac{b}{a} \le \gamma < \frac{\pi}{2}$. If in addition $\gamma \le \arccos\frac{a}{2b}$, than we have $c < b$. So, in this case we have

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, a\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [a\sin\gamma, c) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [c, b) \\ \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b, a) \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases} \tag{4.6}$$

If $a \le b\sqrt{2}$ and $\arccos\frac{a}{2b} \le \gamma \le \arcsin\frac{b}{a}$, than $a\sin\gamma \le b$ and $c \ge b$, so we have two cases: $c \le a$ and $c > b$. By that mean we split the condition $a \le b\sqrt{2}$ into two conditions: $a \le \frac{b\sqrt{5}}{2}$ and $\frac{b\sqrt{5}}{2} \le a \le b\sqrt{2}$. Now let $a \le \frac{b\sqrt{5}}{2}$, $\arccos\frac{a}{2b} \le \gamma < \arccos\frac{b}{2a}$. In this case we have $b\sin\gamma \le a\sin\gamma \le b \le c \le a$, so

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, a\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [a\sin\gamma, b) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [b, c) \\ \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [c, a) \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases} \tag{4.7}$$

When $a \le \frac{b\sqrt{5}}{2}$ and $\arccos\frac{b}{2a} \le \gamma < \arcsin\frac{b}{a}$, we have $b\sin\gamma \le a\sin\gamma \le b \le a \le c$, so

$$\Phi_{a,b}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, b\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \frac{\pi}{2}\right), & \text{if} \quad y \in [b\sin\gamma, a\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{b\sin\gamma}{y}\right) \cup & \\ \cup \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [a\sin\gamma, b) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [b, a) \\ \left(\arccos\frac{b\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right), & \text{if} \quad y \in [a, c) \\ \emptyset, & \text{if} \quad y > c \quad \text{or} \quad y < 0 \end{cases} \tag{4.8}$$

In case of $a \le \frac{b\sqrt{5}}{2}$ and $\arcsin\frac{b}{a} < \gamma < \frac{\pi}{2}$, $\Phi_{a,b}(y)$ is defined by (4.5). Let $\frac{b\sqrt{5}}{2} \le a \le b\sqrt{2}$. It is not difficult to verify that in this case when $\arccos\frac{a}{2b} \le \gamma < \arcsin\frac{b}{a}$, $\Phi_{a,b}(y)$ is defined by the formula (4.7), and when $\arcsin\frac{b}{a} \le \gamma < \arccos\frac{b}{2a}$, $\Phi_{a,b}(y)$ is defined by (4.4), while in the case of $\arccos\frac{b}{2a} \le \gamma < \frac{\pi}{2}$ it's defined by (4.5).

Consider $\gamma \ge \frac{\pi}{2}$ ($c > b$, $c > a$). In this case the function $\frac{b\sin\gamma}{\cos\varphi}$ in the domain $\left(-\frac{\pi}{2}, \arctan\frac{a-b\cos\gamma}{b\sin\gamma}\right)$ increasing, while function $\frac{a\sin\gamma}{\cos(\varphi-\gamma)}$ in the domain

$\left(\arctan\frac{a-b\cos\gamma}{b\sin\gamma},\frac{\pi}{2}\right)$ decreasing, so

$$\Phi_{a,b}(y)=\begin{cases}\left(-\frac{\pi}{2}+\gamma,\frac{\pi}{2}\right), & \text{if} \quad y\in[0,b)\\ \left(\arccos\frac{b\sin\gamma}{y},\frac{\pi}{2}\right), & \text{if} \quad y\in[b,a)\\ \left(\arccos\frac{b\sin\gamma}{y},\gamma-\arccos\frac{a\sin\gamma}{y}\right), & \text{if} \quad y\in[a,c)\\ \emptyset, & \text{if} \quad y>c \quad \text{or} \quad y<0\end{cases} \tag{4.9}$$

Finally, we obtain the following table

Cases	$\Phi_{a,b}(y)$
$a>b\sqrt{2}$ and $\gamma\le\arccos\frac{a}{2b}$, $\quad a\le b\sqrt{2}$ and $\gamma<\arccos\frac{b}{a}$	(4.2)
$a>b\sqrt{2}$ and $\arccos\frac{a}{2b}<\gamma\le\arccos\frac{b}{a}$	(4.3)
$a>b\sqrt{2}$ and $\arccos\frac{b}{a}\le\gamma<\arccos\frac{b}{2a}$, $\frac{b\sqrt{5}}{2}\le a\le b\sqrt{2}$ and $\arcsin\frac{b}{a}\le\gamma<\arccos\frac{b}{2a}$	(4.4)
$a>\frac{b\sqrt{5}}{2}$ and $\arccos\frac{b}{2a}\le\gamma<\frac{\pi}{2}$, $a\le\frac{b\sqrt{5}}{2}$ and $\arcsin\frac{b}{a}<\gamma<\frac{\pi}{2}$	(4.5)
$a\le b\sqrt{2}$ and $\gamma\le\arccos\frac{a}{2b}$	(4.6)
$a\le\frac{b\sqrt{5}}{2}$ and $\arccos\frac{a}{2b}\le\gamma<\arccos\frac{b}{2a}$, $\frac{b\sqrt{5}}{2}\le a\le b\sqrt{2}$ and $\arccos\frac{a}{2b}\le\gamma<\arcsin\frac{b}{a}$	(4.4)
$a\le\frac{b\sqrt{5}}{2}$ and $\arccos\frac{b}{2a}\le\gamma<\arcsin\frac{b}{a}$	(4.8)
$\frac{\pi}{2}\le\gamma<\pi$	(4.9)

Table 1

When segments a and b have no common endpoint, we put them so, that the intersection point of lines containing a and b coincide with the origin, and segment a lie on X-axis. Denote by a' and b' the segments containing a and b, where one of the endpoints of a' and b' coincide with the origin, and another endpoint coincides with the endpoint of a and b correspondingly. Denote $a''=a'\ a, b''=b'\ b$. It is not difficult to verify that

$$I_{[a]\bigcap[b]}(g)=I_{[a']\bigcap[b']}(g)-I_{[a']\bigcap[b'']}(g)-I_{[a'']\bigcap[b]}(g)+I_{[a'']\bigcap[b'']}(g)\,,$$

where I is the characteristic function. Here segments a' and b', a' and b'', a'' and b', a'' and b'' have common endpoint, so we can use the table above.

In fact we obtain that for every convex polygon $\mathbb{D}$ we can find the domains $\Phi_{ij}(y)$ and calculate the integrals in (3.4).

5. CHORD LENGTH DISTRIBUTION FUNCTION FOR RHOMBUS

Consider the special case of (3.4) when $\mathbb{D}$ is a rhombus with side a and angle γ $\left(\gamma \le \frac{\pi}{2}\right)$:

$$\begin{aligned} F(y) = 1 - \tfrac{y}{2a\sin\gamma}\left[\int_{\Phi_{12}(y)} \sin\varphi\,\sin(\gamma-\varphi)\,d\varphi + \int_{\Phi_{14}(y)} \sin\varphi\,\sin(\pi-\gamma-\varphi)\,d\varphi\right] \\ + \tfrac{\sqrt{y^2-a^2\sin^2\gamma}}{2ay}\left[I(a\sin\gamma \le y \le 2a\sin\tfrac{\gamma}{2})\left(a(1-\cos\gamma) - \sqrt{y^2-a^2\sin^2\gamma}\right)\right. \\ +I(a\sin\gamma \le y \le a)\left(a(1-\cos\gamma) + \sqrt{y^2-a^2\sin^2\gamma}\right) \\ \left. +I(a \le y \le 2a\cos\frac{\gamma}{2})\left(a(1+\cos\gamma) - \sqrt{y^2-a^2\sin^2\gamma}\right)\right] - \frac{(a-y)^+}{a}. \end{aligned} \tag{5.1}$$

From Table 1 for $\Phi_{12}(y)$ we have 2 cases: $\gamma \le \frac{\pi}{3}$, and $\frac{\pi}{3} < \gamma \le \frac{\pi}{2}$, which are defined by (4.6), (4.8) correspondingly. Let $\gamma \le \frac{\pi}{3}$. We have

$$\Phi_{12}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, a\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\arccos\frac{a\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [a\sin\gamma, 2a\sin\frac{\gamma}{2}) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{a\sin\gamma}{y}\right) \cup & \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [2a\sin\frac{\gamma}{2}, a) \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases}$$

$\Phi_{14}(y)$ is defined by (4.9):

$$\Phi_{14}(y) = \begin{cases} \left(\frac{\pi}{2}-\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, a) \\ \left(\arccos\frac{a\sin\gamma}{y}, \pi - \gamma - \arccos\frac{a\sin\gamma}{y}\right), & \text{if} \quad y \in [a, 2a\cos\frac{\gamma}{2}) \\ \emptyset, & \text{if} \quad y > 2a\cos\frac{\gamma}{2} \quad \text{or} \quad y < 0 \end{cases}$$

Calculating integrals in (5.1) we obtain

$$F(y) = \begin{cases} 0, & \text{if} \quad y < 0 \\ \frac{y}{2a}\left[1 + \left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right], & \text{if} \quad y \in [0, a\sin\gamma) \\ \frac{y}{2a}\left[1 - \left(\frac{\pi}{2}+\gamma-2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + & \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{y}, & \text{if} \quad y \in [a\sin\gamma, 2a\sin\frac{\gamma}{2}) \\ \frac{y}{4a}\left[3 + \left(2\arcsin\frac{a\sin\gamma}{y} - 3\gamma\right)\cot\gamma\right] + & \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{2y}, & \text{if} \quad y \in [2a\sin\frac{\gamma}{2}, a) \\ 1 + \frac{y}{4a}\left[-1 + \left(\gamma - 2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + & \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{2y} & \text{if} \quad y \in [a, 2a\cos\frac{\gamma}{2}) \\ 1, & \text{if} \quad y \ge 2a\cos\frac{\gamma}{2} \end{cases} \tag{5.2}$$

It is not difficult to verify that function defined by (5.2) is a continuous function. Now let $\frac{\pi}{3} < \gamma \leq \frac{\pi}{2}$. From (4.8) we have

$$\Phi_{12}(y) = \begin{cases} \left(-\frac{\pi}{2}+\gamma, \frac{\pi}{2}\right), & \text{if} \quad y \in [0, a\sin\gamma) \\ \left(-\frac{\pi}{2}+\gamma, -\arccos\frac{a\sin\gamma}{y}\right) \cup \\ \cup \left(\arccos\frac{a\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) \cup \\ \cup \left(\gamma + \arccos\frac{a\sin\gamma}{y}, \frac{\pi}{2}\right) & \text{if} \quad y \in [a\sin\gamma, a) \\ \left(\arccos\frac{a\sin\gamma}{y}, \gamma - \arccos\frac{a\sin\gamma}{y}\right) & \text{if} \quad y \in [a, 2a\sin\frac{\gamma}{2}) \\ \emptyset, & \text{if} \quad y > a \quad \text{or} \quad y < 0 \end{cases}$$

As in previous case $\Phi_{14}(y)$ is defined by (4.9). Finally we have

$$F(y) = \begin{cases} 0, & \text{if} \quad y < 0 \\ \frac{y}{2a}\left[1 + \left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right], & \text{if} \quad y \in [0, a\sin\gamma) \\ \frac{y}{2a}\left[1 - \left(\frac{\pi}{2}+\gamma-2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{y}, & \text{if} \quad y \in [a\sin\gamma, a) \\ 1 - \frac{y}{2a}\left[1 + \left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right] \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{y}, & \text{if} \quad y \in [a, 2a\sin\frac{\gamma}{2}) \\ 1 + \frac{y}{4a}\left[-1 + \left(\gamma - 2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + \\ + \frac{\sqrt{y^2-a^2\sin^2\gamma}}{2y} & \text{if} \quad y \in [2a\sin\frac{\gamma}{2}, 2a\cos\frac{\gamma}{2}) \\ 1, & \text{if} \quad y \geq 2a\cos\frac{\gamma}{2} \end{cases} \tag{5.3}$$

In (5.3) function $F(y)$ is continuous.

Thus, in terms of elementary functions we find the elementary expression of chord length distribution for rhomb. It is defined by formulae (5.2) and (5.3). If we denote by $f(y) = F'(y)$ the chord length density function, then for $\gamma \leq \frac{\pi}{3}$ and for $\frac{\pi}{3} < \gamma \leq \frac{\pi}{2}$ we have correspondingly

$$f(y) = \begin{cases} 0, & \text{if} \quad y < 0 \quad \text{or} \quad y \geq 2a\cos\frac{\gamma}{2} \\ \frac{1}{2a}\left[1 + \left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right], & \text{if} \quad y \in [0, a\sin\gamma) \\ \frac{1}{2a}\left[1 - \left(\frac{\pi}{2}+\gamma-2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + \\ + \frac{a^2\sin^2\gamma + y^2\cos\gamma}{y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y \in [a\sin\gamma, 2a\sin\frac{\gamma}{2}) \\ \frac{1}{4a}\left[3 + \left(2\arcsin\frac{a\sin\gamma}{y} - 3\gamma\right)\cot\gamma\right] + \\ + \frac{a^2\sin^2\gamma - y^2\cos\gamma}{2y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y \in [2a\sin\frac{\gamma}{2}, a) \\ \frac{1}{4a}\left[-1 + \left(\gamma - 2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right] + \\ + \frac{a^2\sin^2\gamma + y^2\cos\gamma}{2y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y \in [a, 2a\cos\frac{\gamma}{2}) \end{cases} \tag{5.4}$$

$$f(y)=\begin{cases}0, & \text{if} \quad y<0 \text{ or } y\geq 2a\cos\frac{\gamma}{2}\\ \frac{1}{2a}\left[1+\left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right], & \text{if} \quad y\in[0,a\sin\gamma)\\ \frac{1}{2a}\left[1-\left(\frac{\pi}{2}+\gamma-2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right]+ & \\ +\frac{a^2\sin^2\gamma+y^2\cos\gamma}{y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y\in[a\sin\gamma,a)\\ -\frac{1}{2a}\left[1+\left(\frac{\pi}{2}-\gamma\right)\cot\gamma\right]+ & \\ +\frac{a\sin^2\gamma}{y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y\in[a,2a\sin\frac{\gamma}{2})\\ \frac{1}{4a}\left[-1+\left(\gamma-2\arcsin\frac{a\sin\gamma}{y}\right)\cot\gamma\right]+ & \\ +\frac{a^2\sin^2\gamma+y^2\cos\gamma}{2y^2\sqrt{y^2-a^2\sin^2\gamma}}, & \text{if} \quad y\in[2a\sin\frac{\gamma}{2},2a\cos\frac{\gamma}{2}),\end{cases} \tag{5.5}$$

The graphs of chord length distribution and density function for rhombus with side $a=1$ and angles $\gamma=\frac{\pi}{6}$ and $\gamma=\frac{5\pi}{12}$ see in Figures 1 and 2.

References

[1] R. Schneider and W. Weil(1990). Stochastic and Integral Geometry, *Springer Verlag*, Berlin.

[2] Ambartzumian, R. (1990) Factorization Calculus and Geometric Probability, *Cambridge University Press*, Cambridge.

[3] Gille, W. (1988). The chord length distribution of parallelepipeds with their limiting cases, *Exp. Techn. Phys.* **36,** 197–208.

[4] Aharonyan, N. and Ohanyan, V. (2005). Chord length distribution functions for polygons, *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences).* **40,** no. 4, 43–56.

[5] Harutyunyan, N. and Ohanyan, V. (2009). Chord length distribution function for regular polygon, *J. Appl. Prob.* **42,** no. 2, 358–366.

[6] Stoyan D. and Stoyan,H. (1994). Fractals, Random Shapes and Point Fields, *John Wiley& Sons*, Chichester, New York, Brisbane, Toronto, Singapore.

[7] Gates J. (1982). Recognition of triangles and quadrilaterals by chord length distribution, *J. Appl. Prob.* **19,** 873–879.

[8] Gates J.(1987). Some properties of chord length distributions, *J. Appl. Prob.* **24** 863–874.

[9] Gille W., Aharonyan N. G. and Haruttyunyan H. S.(2009). Chord length distribution of pentagonal and hexagonal rods: relation to small-angle scattering. In *Journal of Applied Cristallography* 326–328.

[10] Shilov, G. (1984). Mathematical Analysis, (second special course), 2nd edn, *Moscow, Moscow state University.*

[11] R. Sulanke. (1961). Die Verteilung der Sehnenlängen an Ebenen und räumlichen Figuren, *Math. Nachr.*,, **23**, 51 – 74.

[12] V. K. Ohanyan and N. G. Aharonyan (2009). Tomography of bounded convex domains, *Sutra: International Journal of Mathematical Science Education,* **2,** no. 1, 1 - 12,

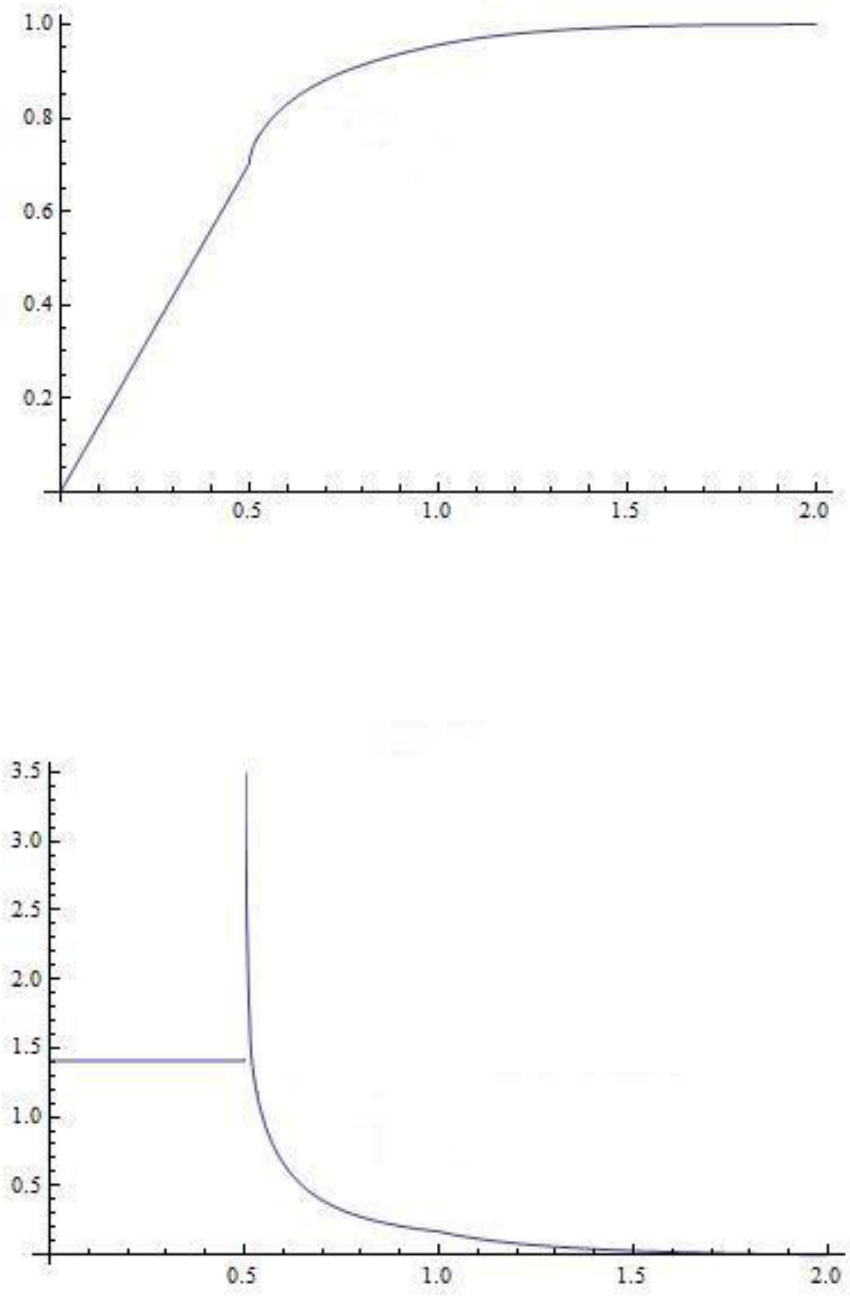


Fig. 1. Chord length distribution and density functions for rhombus with $a = 1$ and $\gamma = \frac{\pi}{6}$.

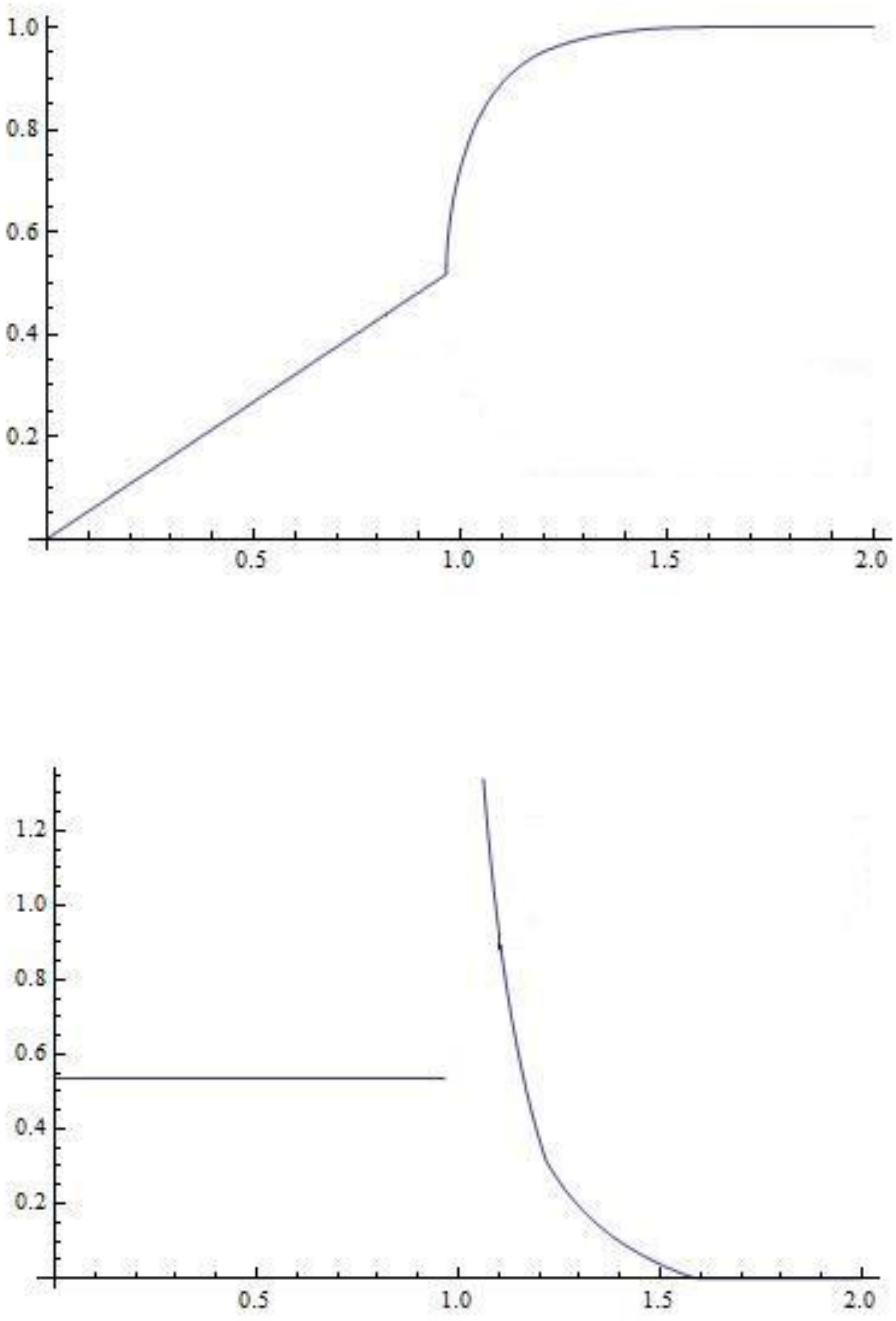


Fig. 2. Chord length distribution and density functions for rhombus with $a = 1$ and $\gamma = \frac{5\pi}{12}$.

Sutra: International Journal of Mathematics Science Education
©Technomathematics Research Foundation
Vol. 4, No. 2, pp. 16 – 26, 2011

TEACHIING ADAPTIVE AND STRATEGIC REASONING THROUGH FORMULA DERIVATION: BEYOND FORMAL SEMIOTICS

ELLIOTT OSTLER
Department of Teacher Education, University of Nebraska at Omaha
6001 Dodge St., Omaha, Nebraska 68182, United States
elliottostler@gmail.com

The paper illustrates how mathematical tasks involving dynamic solution methods such as formula derivations can strengthen mathematical proficiency for secondary level students through exposure to *adaptive* and s*trategic* reasoning.

Keywords: Adaptive Reasoning; Strategic Competence; Procedural Fluency; Solution Dynamics; formula derivation

1. Introduction

Conventional logic in teaching American secondary level students (*secondary level* meaning students comprised of grades 7-12) tells us that mathematics instruction should be both efficient and informative. Logic also suggests that efficiency and information should not be mutually exclusive, and yet American students continue to struggle with adapting their learning beyond the rudiments of their textbooks. This is possibly because the most common instructional method used in secondary mathematics classrooms is highly structured and is based primarily on the semiotic processes so commonly illustrated in mathematics textbooks (Watanabe, 2007). *Semiotic*s in mathematics is simply defined as the use of symbols that are helpful in understanding the processes of thinking, symbolizing, and communicating (Radford, Schubring, & Seeger, 2008). Semiotics in mathematics instruction typically refers to the teaching of mathematics as a process of symbol manipulation through structured algorithms and rigorously defined theorems. The outcome for students in this strategy of teaching and learning is referred to as *procedural fluency*. Procedural fluency is defined as the skill in carrying out procedures flexibly, accurately, efficiently, and appropriately (National Research Council, 2001). Procedurally fluent students ostensibly develop the ability to evaluate and simplify various expressions, solve simple equalities, and represent mathematical relationships in graphical form. However, as the mathematics performance of American secondary students continues to remain low relative to students in other countries, the value of teaching for procedural fluency without identifying the relevance of the underlying mathematical logic remains a matter of debate. The manuscript at hand describes how pedagogy including *Strategic Competence* and *Adaptive Reasoning* can supplement procedural fluency to create a more comprehensive learning experience for secondary mathematics students.

Adaptive Reasoning and Strategic Competence have emerged as two critical components of mathematical proficiency in contemporary teacher training programs in the United

States. *Adaptive Reasoning* is loosely defined as the capacity for logical thinking and the ability to reason and justify why solutions are appropriate within the context of problems that are large in scope, while *Strategic Competence* refers to the ability to formulate suitable mathematical models and select efficient methods for solving problems (National Research Council, 2001). These mathematical strands, along with procedural fluency, are not intended to be independent nor are they intended to be instructed as isolated skills. They are actually highly interdependent and represent different aspects of a more comprehensive kind of mathematics teaching and learning. The examples presented hereafter illustrate how mathematical representations such as formula derivations can be used to develop a balance of procedural fluency, adaptive reasoning, and strategic competence.

The ability to incorporate adaptive reasoning and strategic competence into mathematics instruction is challenging to be certain. Research on teacher lesson planning suggests that reform-based curricula in mathematics presents contemporary challenges that are much different than those lesson development challenges seen within more conventional or traditional curriculums (Superfine, 2008). Thus, the development of instructional tasks that are dynamic in how they present adaptive or strategic tasks is difficult. Nevertheless, adaptive reasoning and strategic competence remain critical elements of a coherent understanding of mathematics. Romberg (2000) argues that with appropriate guidance from teachers, students can build a coherent understanding of mathematics, and that their understanding about how the symbolic processes of mathematics can evolve into increasingly abstract and scientific reasoning. This suggests that supplementing traditional procedural fluency problems with intentional efforts to include adaptive and strategic undertones results in higher levels of comprehension.

The long-term benefits of mathematical exercises, such as those that follow, are also a matter of debate. Many American teachers consider mathematical tasks requiring students to exhibit *adaptive* and *strategic* processing to be vague and inadequate for developing specific (and testable) mathematical skill sets. Mathematical tasks, such as the formula derivations presented herein, have several possible solutions and do not typically emphasize the development any specific kind of procedural fluency. For this reason, they are often viewed as inefficient or even superfluous exercises in the classroom. The counterpoint, however, suggests that vaguely defined tasks are exactly what allow *strategic* and *adaptive* reasoning to occur in a mathematics lesson. A mathematical task having the potential to be completed using numeric, algebraic, geometric, or even calculus-based approaches is especially powerful. The National Council of Teachers of Mathematics (NCTM, 2000) suggests that teachers should use a variety of strategies that allow them to monitor students' capacity to analyze mathematical situations, frame and solve problems and better make sense of the procedures they have been taught. Each of these recommended methods and strategies provides insights into *strategic* and *adaptive* reasoning.

Because the solutions to broader scope mathematical tasks are often dynamic and depend on the mathematical preparedness of the student, the identical mathematical exercises can be most effectively used successively over a period of years to further emphasize the nature and necessity of strategic thinking (Ostler, Grandgenett, & Mitchell, 2008). Students would select and develop appropriate mathematical models and efficient

methods of solving problems based on what their level of mathematical competence allows. These kinds of problems are known as Dynamic Solution Exercises (DSE) and accentuate the need to learn increasingly sophisticated mathematical strategies as a way to more efficiently complete mathematical tasks.

2. Algebraic Formula Derivation (DSE Level 1)

Suppose we want students to present an algebraic solution for deriving the formula for calculating the volume of the frustum of a cone. There are many ways to complete this task. Some are simple while others are more sophisticated and mathematically rigorous. For example, we may begin by constructing a cone for which we already know the formula and cut it to create a frustum. A geometric strategy would be to construct a model that represents the large cone as a frustum and a smaller cone resting on the top (**Fig. 1**). This is a *strategic* model and demands a justification that is more sophisticated than what student might see with traditional semiotic processing. Questioning how and why this method works is what provides the *adaptive* and *strategic* learning.

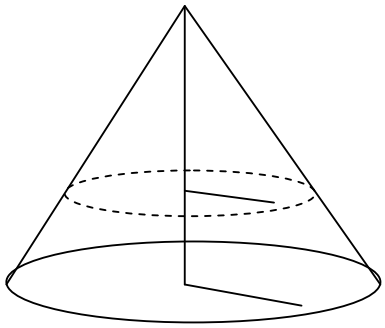


Fig. 1: Cone and frustum

We will assume, base on previous experience, that we know the formula for the volume of a cone. Given this assumption, we then develop algebraic notation for the total volume of the large cone as the sum of the small cone and the frustum. The frustum then must be the difference of the large cone and the small cone. Students would need to justify the algebraic representations of the geometric model as follows:

$$V_{cone} = \frac{\pi}{3} r^2 h \tag{1}$$

$$V_t = V_f + V_c \tag{2}$$

$$V_f = V_t - V_c \tag{3}$$

Defining the variables represents an *adaptive* task because students are not often asked to assign variables in equations they must derive. For the most part, students are simply required to manipulate existing symbols. For this derivation we will label relevant variables as follows: ***k*** (height of the small cone), ***r*** (radius of the base of the small cone), ***h*** (height of the frustum), and ***R*** (radius of the base of the frustum and of the large cone). Substituting the newly defined variables into the equation for a cone (1) and substituting each respective volume equation into equation (3) gives us the following series of equations. Once again the strategic reasoning would need to be emphasized in the classroom.

$$V_f = V_t - V_c \tag{4}$$

$$V_f = \frac{\pi}{3}R^2(h+k) - \frac{\pi}{3}r^2k \tag{5}$$

$$V_f = \frac{\pi}{3}R^2h + \frac{\pi}{3}R^2k - \frac{\pi}{3}r^2k \tag{6}$$

We now have in Eq. (6) a general formula that can be simplified to complete our derivation; however, ***k*** does not formally exist in the frustum, so students must draw from previous algebraic experiences to define ***k*** in terms of the existing variables. The following method illustrates the use of a *proportion*, which requires both adaptive reasoning and strategic competence. Although students may be procedurally proficient at completing proportion tasks, they often experience great difficulty in this kind of logic because the opportunities to adapt algebraic procedures from one task in order to make them appropriate for another are very limited. We solve the proportion for ***k*** in terms of our existing variables and substitute the result shown in Eq. (8) into our frustum Eq. (9). The remainder of the derivation is primarily algebraic; however, note that in Eq. (10) a *rational* form of "one" is necessary to simplify the equation. Students often miss this step.

$$\frac{k}{r} = \frac{h+k}{R} \tag{7}$$

$$k = \frac{hr}{R-r} \tag{8}$$

$$V_f = \frac{\pi}{3}\left[R^2h + R^2\left(\frac{hr}{R-r}\right) - r^2\left(\frac{hr}{R-r}\right)\right] \tag{9}$$

$$V_f = \frac{\pi}{3}\left[\frac{R^2h(R-r)}{R-r} + \left(\frac{R^2hr}{R-r}\right) - \left(\frac{r^2hr}{R-r}\right)\right] \tag{10}$$

$$V_f = \frac{\pi}{3}\left[\frac{R^3h}{R-r} - \frac{R^2rh}{R-r} + \frac{R^2hr}{R-r} - \frac{r^2hr}{R-r}\right] \tag{11}$$

$$V_f = \frac{\pi h}{3}\left[\frac{R^3}{R-r} - \frac{R^2r}{R-r} + \frac{R^2r}{R-r} - \frac{r^3}{R-r}\right] \tag{12}$$

$$V_f = \frac{\pi h}{3}\left[\frac{R^3}{R-r} - \frac{r^3}{R-r}\right] \tag{13}$$

In Eq. (14) below, students are often tempted to consider the task complete. Once again, a simple but necessary factoring procedure requires adaptive reasoning because the need to factor is not immediately apparent. The factoring of the *difference of cubes*, though taught and practiced, is not commonly contextualized in broader problems such as this derivation.

$$V_f = \frac{\pi h}{3}\left[\frac{R^3 - r^3}{R-r}\right] \tag{14}$$

$$V_f = \frac{\pi h}{3}\left[\frac{(R-r)(R^2 + Rr + r^2)}{R-r}\right] \tag{15}$$

$$V_f = \frac{\pi h}{3}(R^2 + Rr + r^2) \tag{16}$$

The student outcome of this task is exposure to a simple algebraic derivation that makes complex assumptions about the kinds of adaptive reasoning abilities that students possess. The point of the example is that procedural fluency in *solving proportions*, *renaming fractions*, or *factoring polynomials* does not necessarily mean students are able to adapt the procedures for use in tasks such as the previous formula derivation. Special instructional emphases on how these adaptive and strategic processes are carried out is necessary in order for students to become proficient in these kinds of skills. Additionally, it is worthwhile to compare the value of the algebraic method used above with a more sophisticated procedure in order to develop advanced *strategic* skills. Students must be

exposed to *strategic* thinking to learn to model tasks such as the frustum so that appropriate methods can be selected and carried out in similar tasks. Strategic thinking is often the result of witnessing a kind of problem, such as a formula derivation, completed using a number of different strategies and then demonstrating the ability to select and use the most efficient method for similar kinds of tasks. This is a form of *Solution Dynamics* (i.e. the solution process is a *dynamic* process depending on the preparation and ability of the student) that allows for the completion of a single problem at multiple levels. An illustration of this would be to complete the frustum formula derivation in a more sophisticated way. One possible calculus-based solution for deriving the formula for the volume of a frustum is illustrated below.

3. Calculus Formula Derivation (DSE Level 2)

In this derivation, the process of defining the model and the variables of radii and height for the frustum is more formal than what was done using the algebraic method. We use ordered pairs to define distances rather than assigning variable lengths to arbitrary symbols as was done in the algebraic example. This makes the variables more tangible to students, but also makes adapting certain algebraic processes more difficult for them. Instruction in how variables are assigned is important and must include opportunities to explore the corroborating algebraic relationships since the lengths of different parts of the frustum are derived from coordinate values within the ordered pairs, and the shape of the frustum itself is defined by the equation of a general linear equation.

The process used in the following example is calculus-based and is difficult for students from a strategic standpoint only in that they must recognize how to translate a single-integration area model to compute volume by using a disk-method. This is done using a general linear equation and rotating it around an axis. The remainder of the problem is procedural and largely algebraic, but it is still worth comparing to the previous method for strategic purposes. In particular, the strategy does not use the differences of two cones to geometrically define the frustum. Instead, a simple linear equation is used to define the frustum. In this example the frustum starts with center point of the top circular surface at the point (0,0) and with the bottom circular surface terminating at the point (***h***,0). A general linear equation using points at the end of each respective radius, the top defined by the point (0,***r***) and the bottom defined by the point (***h***,***R***), determine the slant of the frustum as illustrated in **Fig. 2**.

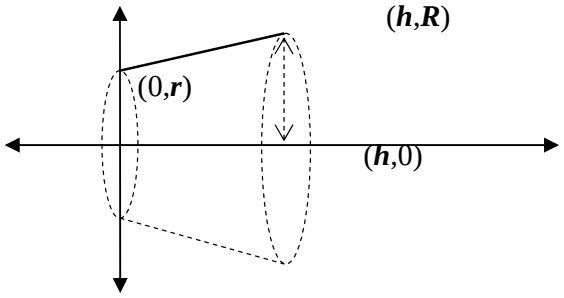


Fig. 2: Frustum with top and bottom bases at (0,0) and (h,0) respectively; and small and large radii of r and R respectively.

Note that the equation of the line defined by the given points must be derived using slope-*y* intercept equation. The slope is then used to define general linear equation, which can be rotated around the x-axis to create a series of disks. This is both an adaptive and a strategic skill. It is helpful to have students justify the parts of the resulting differential equation and describe the process of adding a series of disks whose radii are successively determined by the height of the liner function. The purpose is to demonstrate the adaptive reasoning needed for describing how discrete calculations are transferred to continuous methods.

$$m = \frac{R-r}{h-0} = \frac{R-r}{h} \tag{17}$$

$$f(x) = \left(\frac{R-r}{h}\right)x + r \tag{18}$$

Eq. (18), the function that will represent the radius of each of the disks, is going to be substituted in for the radius variable (***r***) in Eq. (19).

$$\int_0^h \pi r^2 dx \tag{19}$$

$$\int_0^h \pi\left[f(x)\right]^2 dx \tag{20}$$

$$\int_0^h \pi\left[\left(\frac{R-r}{h}\right)x + r\right]^2 dx \tag{21}$$

$$\pi\int_0^h \left[\left(\frac{R-r}{h}\right)^2 x^2 + 2r\left(\frac{R-r}{h}\right)x + r^2\right]dx \tag{22}$$

A few basic rules of integration are applied and the antiderivative calculated before the variable ***h*** is substituted into Eq. (24). The resulting algebra yields Eq. (29), which is the same volume formula for the frustum as was derived previously.

$$\pi\left[\left(\frac{R-r}{h}\right)^2\frac{x^3}{3}+2r\left(\frac{R-r}{h}\right)\frac{x^2}{2}+r^2x\right]_0^h \tag{23}$$

$$\pi\left[\left(\frac{R^2-2Rr+r^2}{h^2}\right)\frac{h^3}{3}+\left(\frac{2Rr-2r^2}{h}\right)\frac{h^2}{2}+r^2h\right]-\pi[0] \tag{24}$$

$$\pi\left[\left(\frac{R^2-2Rr+r^2}{1}\right)\frac{h}{3}+\left(\frac{Rr-r^2}{1}\right)\frac{h}{1}+r^2h\right] \tag{25}$$

$$\pi\left[\left(\frac{R^2-2Rr+r^2}{1}\right)\frac{h}{3}+\left(\frac{Rr-r^2}{1}\right)\frac{(3)h}{(3)1}+\frac{(3)r^2h}{(3)}\right] \tag{26}$$

$$\frac{\pi h}{3}\left[R^2-2Rr+r^2+\left(\frac{Rr-r^2}{1}\right)\frac{(3)}{1}+\frac{(3)r^2}{1}\right] \tag{27}$$

$$\frac{\pi h}{3}\left[R^2-2Rr+r^2+3Rr-3r^2+3r^2\right] \tag{28}$$

$$V_f=\frac{\pi h}{3}\left(R^2+Rr+r^2\right) \tag{29}$$

Alternative methods for completing this task (using calculus) may initially seem redundant but they are important for the development of both procedural fluency and strategic competence. Students who are not given the opportunity to see how dynamic processes develop are not able to adapt their thinking to other, more complex mathematical situations.

4. A Comparative Example of Strategic Competence

In the introduction of this manuscript, strategic competence was roughly defined as the ability to formulate suitable mathematical models and select efficient methods for solving

problems. Consider the following derivation. Suppose we want to derive the formula for calculating the area of a circle. Logic, of course, suggests that we simply use a known reference; however, if the task is to demonstrate the ability to derive the equation, then the *strategic* task is to select the most efficient task using available information. There are many forms of algebraic, geometric, trigonometric, and calculus based derivations for completing the circle area formula, but extending our strategic abilities for the purposes of efficiency might look like the following comparison of two methods. In the first method, the circle is graphed on a coordinate plane with the center at (0,0) and with a radius of ***r***. We then solve for *y*, representing it as a function and integrate the resulting form from 0 to ***r***. The resulting formula represents only one fourth of the circle. The derivation is illustrated below:

$$x^2 + y^2 = r^2 \tag{30}$$

$$y = \sqrt{r^2 - x^2} \tag{31}$$

$$\int_0^r \sqrt{r^2 - x^2}\ dx \tag{32}$$

$$\frac{u}{2}\sqrt{a^2 - u^2} + \frac{a^2}{2}\sin^{-1}\frac{u}{a} \tag{33}$$

$$\frac{x}{2}\sqrt{r^2 - x^2} + \frac{r^2}{2}\sin^{-1}\frac{x}{r}\Big|_0^r \tag{34}$$

$$\left[\frac{r}{2}\sqrt{r^2 - r^2} + \frac{r^2}{2}\sin^{-1}\frac{r}{r}\right] - \left[\frac{0}{2}\sqrt{r^2 - 0^2} + \frac{r^2}{2}\sin^{-1}\frac{0}{r}\right] \tag{35}$$

$$\left[\frac{r}{2}(0) + \frac{r^2}{2}\sin^{-1}1\right] - \left[\frac{0}{2}(r) + \frac{r^2}{2}\sin^{-1}0\right] \tag{36}$$

$$\left[0 + \frac{r^2}{2}\cdot\frac{\pi}{2}\right] - \left[0 + \frac{r^2}{2}(0)\right] \tag{37}$$

$$\frac{\pi r^2}{4} \tag{38}$$

Now let us consider a different more simple and efficient technique using the known circumference formula of a circle. Visualize cutting roughly circular strips from the circle that look like thin circular bands. If the bands are cut, they become roughly rectangular. The lengths of the rectangular strips are analogous to the circumference of the circle and the widths of the strips are the differences in ***r*** from the inside of each strip to the outside. Adding all of the strips together by integrating from the center of the circle to the edge of the circle at radius ***r*** will result in the area formula we seek. Well conceived, this is simply a matter of integrating the circumference over the length of the radius.

$$\int_0^r 2\pi r \; dr \tag{39}$$

$$2\pi \int_0^r r \; dr \tag{40}$$

$$2\pi \left(\frac{r^2}{2} \right) \Big|_0^r \tag{41}$$

$$\pi r^2 \tag{42}$$

The derivation is completed in a few simple steps, which represents strategic insight. This method can also be easily adapted to a three dimensional model where the surface area of a sphere can be integrated to derive the formula for the volume. Repeated exposure to such derivations not only provides procedural fluency for different maths, it also lends credibility to strategic and adaptive reasoning for fundamentally equivalent mathematical tasks.

5. Conclusion

The focus on using formulas has become so common in secondary mathematics instruction that we have underestimated the value of contextualizing the processes that allows them to exist. Perhaps the derivations presented previously illustrate an over simplified approach to a very complex problem, but unless teachers truly consider the innate limitations of the way we teach procedural fluency, we will perpetuate the belief that procedures alone represent what constitutes good mathematical learning and what makes good mathematicians.

References

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Ostler, E., Grandgenett, N., & Mitchell, C. (2008). *Rethinking mathematics assessment: Some reflections on solution dynamics as a way to enhance quality indicators*. The Mathematics Educator. Vol. 18. No. 2, 33-39.

Radford, L., Schubring, G., & Seeger, F. (2008). *Semiotics in Mathematics Education: Epistemology, History, Classroom, and Culture*. Sense Publishers: Rotterdam/Taipei.

Romberg, T. (2000). *Changing the teacher and learning of mathematics*. Australian Mathematics Teacher, Vol. 56, No. 4, 6-9.

Superfine, A. C., (2008). *Planning for mathematics instruction: A model of experienced teachers' planning processes in the context of a reform mathematics curriculum*. The Mathematics Educator. Vol. 18, No. 2, 11-22.

Watanabe, T. (2007). *In pursuit of a focused and coherent school mathematics curriculum*. The Mathematics Educator. Vol. 17, No. 1, 2-6.

Sutra: International Journal of Mathematics Science Education
©Technomathematics Research Foundation
Vol. 4, No. 2, pp. 27 – 29, 2011

Solving Cubics by Division Method

Raghavendra G. Kulkarni
Bangalore, India

Abstract

We present a division method for factoring and solving cubics. The method derives formulas in a logical sequence rather than in an intuitive way (as given in the Viesta's substitution method).

1 Introduction

It is well-known that Cardano (1539) first published his solution to the gen- eral cubic equation in his book "The great art, or, the rules of algebra" using the concept of complex numbers, which was unknown at that time. However, history pages reveal that Scipione del Ferro (1515) had solved the cubic in a mathematical contest, but didn't publish his method; and it seems Tartaglia (1535) revealed his solution to Cardano, who got it published in his name [1, 2].

In this paper, we describe a division method for factoring and subsequently solving the cubic equation, which relies on derivation of formulas in a logical fashion instead of intuitive or empirical way, the Viesta's substitution proposes [2].

2 The division method

Consider the following reduced cubic polynomial:

$$p(x) = x^3 + a_1 x + a_0 \tag{1}$$

for factoring and solving by the proposed division method, where the coefficients, a_0 and a_1, are real. Let us consider the quadratic polynomial, $q(x)$, as given below:

$$q(x) = x^2 + (m+n)x + k \tag{2}$$

where m, n, and k, are unknowns to be determined. Dividing cubic (1) by the quadratic (2), we obtain the quotient as:

$$L(x) = x - (m+n) \tag{3}$$

and the remainder as:

$$R(x) = [a_1 - k + (m+n)^2]x + [a_0 + k(m+n)] \tag{4}$$

If the quadratic (2) has to be a factor of cubic (1), then the remainder, $R(x)$, has to be zero for all values of x, which means each of the coefficients of x and x^0 in $R(x)$ has to be zero, as shown below.

$$a_1 - k + (m+n)^2 = 0 \tag{5}$$

$$a_0 + k(m+n) = 0 \tag{6}$$

Using (6) we eliminate k from the expression (5), resulting in a cubic equation in $(m+n)$ as shown below.

$$(m+n)^3 + a_1(m+n) + a_0 = 0 \tag{7}$$

Expanding the above expression and rearranging, we obtain:

$$m^3 + n^3 + (m+n)(a_1 + 3mn) + a_0 = 0 \tag{8}$$

Since we have only one equation [(8)] in two unknowns (m and n), we require one more equation to determine the unknowns. Therefore we set n as:

$$n = \frac{-a_1}{3m} \tag{9}$$

and use it in (8), which converts (8) into a quadratic equation in m^3 as shown below.

$$m^6 + a_0 m^3 - \frac{a_1^3}{27} = 0 \tag{10}$$

Solving (10), we obtain two values of m^3 as:

$$m^3 = \left(\frac{-a_0}{2}\right) \pm \frac{1}{2}\sqrt{a_0^2 + \frac{4a_1^3}{27}} \tag{11}$$

Taking the cube-root of (11), we determine m as:

$$m = \left[\left(\frac{-a_0}{2}\right) \pm \frac{1}{2}\sqrt{a_0^2 + \frac{4a_1^3}{27}}\right]^{1/3} \tag{12}$$

Since m is already determined, the remaining unknowns [in $q(x)$], n and k, are determined from (9) and (6) respectively. Thus all unknowns in the quadratic polynomial, $q(x)$, are determined satisfying the condition that the remainder, $R(x) = 0$, for all values of x. Hence the cubic polynomial, $p(x)$, can be expressed as product of its linear [quotient $L(x)$] and quadratic [divisor $q(x)$] factors as shown below.

$$p(x) = L(x)q(x) \tag{13}$$

Equating each of the factors in the above cubic to zero and solving, we obtain all the three roots of given cubic (1) as shown below.

$$x_1 = m - \frac{a_1}{3m} \tag{14}$$

$$x_2 = \frac{1}{2}\left\{\left(m - \frac{a_1}{3m}\right)\left[-1 + \sqrt{1 + \frac{4a_0}{\left(m - \frac{a_1}{3m}\right)^3}}\right]\right\} \tag{15}$$

$$x_3 = \frac{1}{2}\left\{\left(m - \frac{a_1}{3m}\right)\left[-1 - \sqrt{1 + \frac{4a_0}{\left(m - \frac{a_1}{3m}\right)^3}}\right]\right\} \tag{16}$$

Let us solve one numerical example using the proposed method. Consider the cubic equation: $x^3 - 2x + 4 = 0$; first, one value of m^3 is determined from (12) as: -0.0755, and taking its cube-root m is determined as: -0.42265. From (17) one root of cubic is found out as: -2. From (18) and (19), the other two roots are determined as: $1 - 1i$, and $1 + 1i$, where $i = \sqrt{-1}$.

3 Conclusions

A division method for solving cubic equation is described, wherein the formulas are derived in a logical fashion to obtain the three roots of given cubic.

Acknowledgements The author thanks the management of Bharat Electronics Ltd., Bangalore for supporting this work.

References

[1] G. Birkhoff and S. MacLane, "A survey of modern algebra", 5th edition, A. K. Peters, 1997.

[2] Eric W. Weisstein, "Cubic Equation", From MathWorld-A Wolfram Web Resource, http://mathworld.wolfram.com/CubicEquation.html.

Sutra: International Journal of Mathematics Science Education
©Technomathematics Research Foundation
Vol. 4, No. 2, pp. 30 – 31, 2011

CORRECTION TO "A NEW METHOD FOR SOLVING QUARTICS"

KENNETH V. CARTWRIGHT

School of Mathematics, Physics and Technology,
College of The Bahamas,
P.O. Box N4912,
Nassau, N.P., Bahamas,
kvcartwright@yahoo.com
http://sites.google.com/site/kvcartwright/

An error in Eq. (14) of the paper "A new method for solving quartics" (http://www.tmrfindia.org/sutra/v2i23.pdf) is identified and corrected.

Keywords: Quartic equation, cubic equation, polynomial decomposition.

Kulkarni (2009) has recently introduced a new method of solving quartics. Unfortunately, a key equation in the method is typed incorrectly. If a reader tried to apply the method without first checking all equations, the correct roots of the quartic will not be gotten, as this author found out the hard way. The error is in Eq. (14): specifically, the a_1a_2 term should in fact be a_1a_3.

To see this, recall that Eq. (14) is derived from Eq. (12) which states that

$$p\{[p+a_2-a_3^2/4]^2-4a_0\}=\{(a_3/2)(p+a_2-a_3^2/4)-a_1\}^2. \qquad (12)$$

Also, Kulkarni (2009) defined $F_1=a_2-a_3^2/4$ and $p_1=p+F_1$ in Eq. (13). Substituting (13) into (12) gives

$$\begin{aligned}
&(p_1-F_1)\left(p_1^2-4a_0\right)=\left\{\left(a_3/2\right)p_1-a_1\right\}^2\\
&\text{or } p_1^3-p_1 4a_0-F_1p_1^2+F_1 4a_0=\left(a_3^2/4\right)p_1^2-a_1a_3p_1+a_1^2\\
&\text{or } p_1^3-\left(F_1+a_3^2/4\right)p_1^2+\left(a_1a_3-4a_0\right)p_1+F_1 4a_0-a_1^2=0\\
&\text{or } p_1^3-a_2p_1^2+\left(a_1a_3-4a_0\right)p_1+4a_0F_1-a_1^2=0. \qquad (14)
\end{aligned}$$

Thus Eq. (14) should contain the term a_1a_3 and not the term a_1a_2 as given in (14) of the original paper.

Fortunately, the cubic equation given in the numerical example in Kulkarni (2009) is correct, even though it was obtained from Eq. (14). This suggests then that the error maybe simply typographical.

Reference

Kulkarni, R. G. (2009): A new method for solving quartics, *Sutra: International Journal of Mathematical Science Education*, **2**, pp. 24-26.

The Wayback Machine - https://web.archive.org/web/20210309184539/http://www.tmrfindia.org/sutra.…

Go to: Content | Footer

Sutra

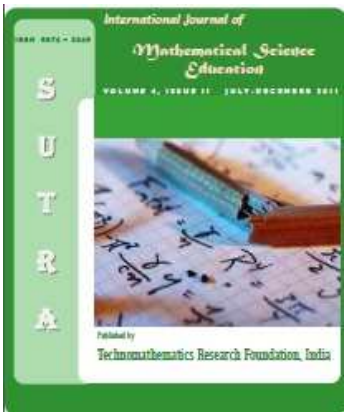
International Journal of Mathematical Science Education

Published Issues

ISSN 0974 – 3340

The journal has stopped publishing new issues from 2012...

CONTENTS
(Vol. 4, No. 2, 2011)



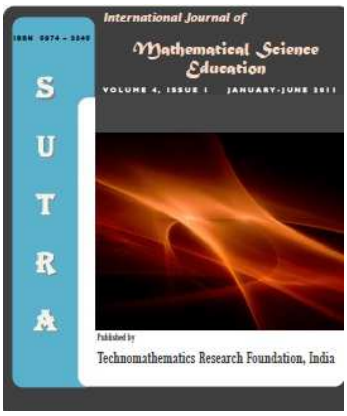
Cover Page

CHORD LENGTH DISTRIBUTION FUNCTION FOR CONVEX POLYGONS
H. S. HARUTYUNYAN AND V. K. OHANYAN

TEACHIING ADAPTIVE AND STRATEGIC REASONING THROUGH FORMULA DERIVATION: BEYOND FORMAL SEMIOTICS
ELLIOTT OSTLER

SOLVING CUBICS BY DIVISION METHOD
RAGHAVENDRA G. KULKARNI

CORRECTION TO "A NEW METHOD FOR SOLVING QUARTICS"
KENNETH V. CARTWRIGHT



CONTENTS
(Vol. 4, No. 1, 2011)

Cover Page

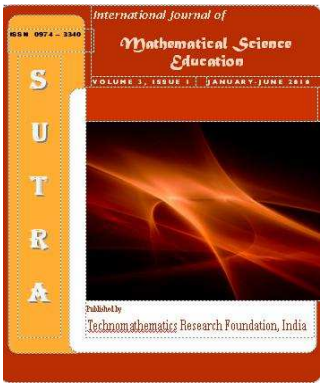
NECESSARY AND SUFFICIENT OPTIMALITY CONDITIONS FOR NONLINEAR FUZZY OPTIMIZATION PROBLEM
V. D. PATHAK · U. M. PIRZADA

SUBCLASSES OF P-VALENT STARLIKE FUNCTIONS DEFINED BY USING CERTAIN FRACTIONAL DERIVATIVE OPERATOR
SOMIA MUFTAH AMSHERI & VALENTINA ZHARKOVA

ANALYSIS OF THE ROLE OF ENTROPIES AND ASSIGMENT OF RANKS TO THE FEATURES IN GENRE DISCRIMINATION
HEMLATA PANDEA & H. S. DHAMIB

IDENTIFYING AND SEQUENCING OF THE ELEMENTARY CONCEPTS OF MEASUREMENT OF MASS
S K. SAMSUL ALAM

THE CENTROID OF SOME GENERALIZED PEDAL CONFIGURATIONS
SZILÁRD ANDRÁS



CONTENTS
(Vol. 3, No. 1, 2010)

Cover Page and Editorial

1. MEASURE OF PLANES INTERSECTING A CONVEX BODY
Rafik Aramyan

2. THE SUM OF ONE SERIES Euler and Bernoulli summation methods of inverted squares series
Frantisek Mosna

3. BOUNDED INVERSE RELATIONAL FUNCTION
Chandan Kumar

4. NUMBERS : RATIONAL, IRRATIONAL OR TRANSCENDENTAL?
Kallol Paul and Dipanwita Paul Ghosh

5. A MIXED QUADRATURE RULE FOR NUMERICAL INTEGRATION OF ANALYTIC FUNCTIONS BY USING BIRKHOFF YOUNG AND BOOLS QUADRATURE
Sanjit Ku. Mohanty and Rajani.B. Dash

6. ETHNOMATHEMATICS: A Mathematical Game in Hausa Culture
Mohammed Waziri Yusuf, Ibrahim Saidu, and Aisha Halliru

7. A COEFFICIENT INEQUALITY FOR A SUB-CLASS OF ANALYTIC FUNCTIONS
B. S. Mehrok, Deepak Gupta and Gagandeep Singh

CONTENTS
(Vol. 2, No. 2, 2009)

Cover Page and Editorial
R. Akerkar

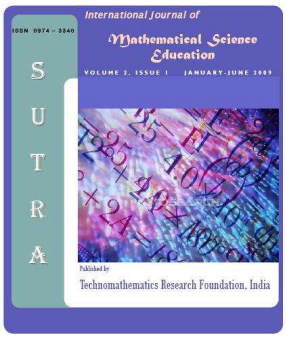
1. **Some Applications of Eulerian Graphs**.
Abdul Samad Ismail , Roslan Hasni and K. G. Subramanian

2. **Engaging Stem Students in Research and Presentation Activities**.
Rohitha Goonatilake and Rafic A. Bachnak

3. **A New Method for Solving Quartics**.
Raghavendra G. Kulkarni

4. **A Number Theoretic Identity**.
Amiya Bhusan Sarkar

CONTENTS
(Vol. 2, No. 1, 2009)



Cover Page and Editorial
R. Akerkar

1. Tomography of Bounded Convex Domains.
V. K. Ohanyan and N. G. Aharonyan

2. Deceptive Patterns in Mathematics.
V. G. Tikekar

3. Harnessing Geometric Manipulatives as a Revitalization Strategy for Mathematics Education in Nigeria
Felix Ehizibue Aburime

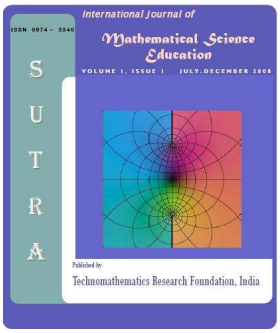
4. Mathematical Puzzle.

5. Solutions to Earlier Puzzles.

6. Interesting Facts about *e*

7. Useful Mathematical Symbols.

CONTENTS
(Vol. 1, No. 1, 2008)



Cover Page and Editorial
Rupali Akerkar

1. Mathematics Thrillerium.
   L. Radhakrishna

2. Accurate Solution Estimate and Asymptotic Behavior of Nonlinear Discrete System.
   Trupti P Shah and Raju K George

3. On Divisibility Tests And The Curriculum Dilemma
   Yaakov Nahir.

4. Class-room Notes.
   R. Akerkar

5. Mathematical Puzzle.

Unless stated explicitly and in conformance to the legal Disclaimer of TMRF Kolhapur, the copyright for the SUTRA as an online journal is with the publisher of this site. The author of individual items (subsuming any type of computer-represented files containing articles, software demos, videos, etc) reserves all proprietary rights such as patent rights and the right to use all or part of the item in future works of their own such as lectures, press releases, and reviews of books. Copying of items, in particular articles, and SUTRA volumes is permitted only for private and academic purposes. Copying or use for commercial purposes is forbidden unless an explicit permission is acquired from the copyright owner. Re-publication of a SUTRA volume(s) or of an individual item inside a journal requires permission by the Chairman, TMRF.
Mirroring of the TMRF SUTRA web site, or parts of it, is prohibited. The label 'SUTRA: International Journal of Mathematical Science Education' and the 'TMRF logo' are owned by the publisher of this site.

Read Submission Guidelines before sending the manuscript to Sutra »

Sutra ...

Welcome!

This is an online free-access journal of Technomathematics Research Foundation. Visit TMRF!

Navigate:

- Aims & Scope
- Editorial Board
- Current Issue
- Submission Guidelines
- Copyright Form
- Referee Form

- Special Issue Proposal

Latest News:

March 1, 2012
There will be no new issues of Sutra from 2012

Links:

- **CFP for Special Issue**
- TMRF Home
- Upcoming Conferences

Contact us | Back to top
© 2008-15 **TMRF Kolhapur**